

Diferencias regionales en las estructuras de ingresos laborales en Colombia *

Karen Chavez

Tesis de Maestría dirigida por:
Juan Miguel Gallego y Carlos Eduardo Sepúlveda
Universidad del Rosario

Palabras clave: Mercado laboral. Desigualdad de ingresos. Diferencias regionales. Oaxaca-Blinder. *Reweighting Regression*.

JEL Classification J31 .

*Este documento es el resultado de una investigación conducida bajo la guía de Juan Miguel Gallego y Carlos Eduardo Sepúlveda, a quienes agradezco su tiempo, asesoría y enseñanzas. También agradezco a Javier Pérez y Juan Carlos Guataqui por sus comentarios como jurados.

1. Introducción

Las diferencias entre mercados laborales regionales se explican desde la teoría neoclásica por diferencias en *amenities* de las ciudades y características de la fuerza laboral; pero dada la persistencia observada en los datos de estas diferencias, surge la teoría de segmentación que, aunque a diferencia de la teoría neoclásica admite diferencias en las estructuras de determinación de ingresos laborales, carece de una base teórica sólida. Es por ello que hasta el momento, tampoco existe un consenso sobre cómo abordar empíricamente la diferenciación de mercados laborales regionales.

En Colombia el análisis desde la perspectiva regional es relevante porque la diferencia entre ciudades no solo está en indicadores de mercado laboral, pues el crecimiento económico y demográfico también varían entre regiones y se ha encontrado evidencia de la no convergencia del PIB entre regiones colombianas (Galvis, 2013; Galvis and Meisel, 2012; Bonilla, 2008); y esto es importante porque se observa persistencia en la desigualdad y establecimiento de trampas de pobreza regionales (Galvis and Meisel, 2012), las cuales pueden contrarrestarse reduciendo las disparidades regionales y el círculo vicioso en el que se encuentran las regiones más deprimidas (Elhorst, J. Paul, 2000). Aún más, estas diferencias están relacionadas con diferencias en la movilidad social (Galvis and Meisel, 2016), y desde la teoría de convergencia regional se ha encontrado que la disparidad está relacionada negativamente con el bienestar agregado (Bérnabou, 1996).

De esta manera, entender las diferencias en los mercados laborales es fundamental para establecer si los análisis y el diseño de políticas encaminados a reducir las brechas de ingresos entre regiones, a partir de un enfoque de ingresos laborales, deben conducirse desde la perspectiva regional o la nacional. Así también se puede contribuir a explicar el poco impacto de las políticas de desarrollo productivo (Nupia, 2014; Eslava et al., 2014; Medina and Saavedra, 2012).

Además, desde el punto de vista de política pública existen dos preguntas claves en el diseño de instrumentos que busquen reducir las disparidades de ingresos laborales entre regiones: ¿son diferenciados los mercados laborales? Y si lo son ¿en qué se diferencian? Acerca de la primera pregunta, estudios previos para Colombia han documentado la heterogeneidad de

los mercados laborales por ciudades, identificando amplios rangos de variación en variables como ingresos laborales, informalidad, tasa global de participación, tasa de ocupación, tasa de subempleo y tasa de desempleo (Arango, 2011); y aunque existen varios estudios que se han concentrado en el análisis de las diferencias de ingresos laborales y exponen la existencia y persistencia temporal de estos diferenciales, aún no se han establecido las causas de esta diferenciación.

De esta manera, la búsqueda de explicaciones a las diferencias observadas son un tema aún pendiente en la literatura; y es por ello que el objetivo de este trabajo es presentar una primera aproximación en la exploración de respuestas a la segunda pregunta y determinar si las diferencias de ingresos laborales entre ciudades se explican por las diferencias en las características de los trabajadores y/o por las diferencias en la forma en la que se remuneran estas características según la estructura de ingresos laborales de cada ciudad.

Siguiendo a Mincer (1958), se asume que los ingresos laborales están determinados por las características de los trabajadores, como educación, experiencia y género; y por la forma como las ciudades remuneran esas características según su estructura de ingresos. Para ello se establece una relación estadística entre el ingreso de un trabajador i ubicado en la ciudad g y sus características (X_i, ε_i) , a través de la función m_g ; por lo tanto, la ecuación que describe el ingreso de un trabajador tiene la siguiente forma $y_{gi} = m_g(X_i, \varepsilon_i)$. Y desde el punto de vista econométrico se buscan aislar las diferencias en las funciones m_g de la diferencia total, lo que corresponde a identificar las diferencias en la forma como se relacionan las características y los ingresos laborales; y que intuitivamente representan diferencias en los retornos a esas características.

Para ello, este trabajo propone como metodología el uso una técnica de descomposición del tipo Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1974), porque estas permiten cuantificar y descomponer las diferencias de ingresos laborales entre dos ciudades (o dos grupos de ciudades) $g = A$ y B , en una parte atribuida a diferencias en características de los trabajadores (efecto composición) y otra parte atribuida a diferencia en los retornos a esas características (efecto estructura). Además, se sigue la discusión del problema de identificación de los términos de la diferencia presentada en Fortin et al. (2011), quienes hacen una analogía entre los métodos de descomposición y los métodos de evaluación de impacto; aclarando que la analogía men-

cionada permite validar el presente ejercicio desde el punto de vista econométrico, aunque el efecto estimado no corresponde a un efecto causal.

Se elige el método *Reweighting Regression* (Fortin et al., 1996), que dentro de la familia de métodos de evaluación de impacto corresponde a una técnica de emparejamiento. Este método es muy informativo porque se logran caracterizar las diferencias a lo largo de toda la distribución, lo cual es particularmente importante cuando se tienen distorsiones como el salario mínimo que influyen los estadísticos calculados sobre la distribución de ingresos laborales. Además, con relación a otros métodos de descomposición del tipo Oaxaca-Blinder, es el que menos supuestos impone sobre la estructura de los datos; pues corresponde a un método semi-paramétrico que no asume una forma lineal para m_g , y para realizar la descomposición en estadísticos diferentes a la media no requiere el establecimiento de supuestos adicionales, que sí están implícitos en otros estadísticos (en el ejemplo más sencillo, al descomponer la resta de la varianza entre dos variables se tienen términos cruzados que son cero solo si se realizan supuestos adicionales).

Debido a que los análisis conducidos se deben realizar entre pares de ciudades (o grupos de ciudades), como ciudad de referencia se elige a Bogotá y como grupos de comparación se eligen:

- Las 12 ciudades principales agrupadas.
- Otras ciudades principales: Medellín, Cali, Barranquilla y Bucaramanga. Pues entre estas cinco ciudades se distinguen dos grupos: aquellas con ingresos superiores a la media (Bogotá, Medellín y Bucaramanga) y aquellas con ingresos inferiores a la media (Cali y Barranquilla).

Acerca de lo anterior, es preciso mencionar que los enfoques propuestos no son únicos y se deja abierta la posibilidad para explorar las diferencias entre mercados regionales para distintos grupos de ciudades; en esta tesis de maestría la elección de estos responde a la importancia relativa de los mercados laborales, y con ellos se lleva a cabo un ejercicio de análisis de diferencias entre mercados regionales bajo la metodología descrita.

Adicionalmente, se argumenta la elección de Bogotá como ciudad de referencia porque es la ciudad con mayor número de habitantes y tiene las mayores tasas de crecimiento poblacional

tanto nominal como absoluto -actuales y proyecciones para los próximos 20 años (DNP, 2012); y en términos de mercado laboral, es la ciudad con mayor número de ocupados y mayores ingresos laborales.

En los casos mencionados se realizan las estimaciones para el total de la muestra; y con la muestra dividida teniendo en cuenta dos tipos de diferenciación de los trabajadores: formal-informal y asalariado-cuenta propia. Lo anterior, con el fin de corregir posibles sesgos en la estimación y en base a los resultados de trabajos previos que han encontrado diferencias significativas entre estos grupos de trabajadores (Badillo et al., 2009; García et al., 2009).

De forma general, los resultados obtenidos al realizar las descomposiciones de las diferencias de ingresos laborales muestran que las diferencias de ingresos laborales entre Bogotá y otras ciudades no se explican completamente por las diferencias en las características de los trabajadores. Adicionalmente, se precisa que los resultados son heterogéneos a lo largo de toda la distribución y entre los distintos grupos de trabajadores.

De forma particular, los resultados obtenidos al comparar a Bogotá con el grupo de las otras 12 ciudades muestran que las diferencias en estructura, en la mayoría de los casos, explican algo más del 50 % de las diferencias observadas en ingresos laborales, y que para los trabajadores cuenta propia las diferencias en estructura de ingresos laborales son mayores a medida que se avanza en los percentiles de la distribución de ingresos laborales.

Y al comparar Bogotá con las otras cuatro ciudades principales, se evidencia que Medellín es la ciudad que cuenta con la estructura de ingresos laborales más cercana a Bogotá; mientras que Bucaramanga es la que menos similitudes en estructura tiene con Bogotá, aunque sus niveles de ingresos no son los que más difieren. Y por su parte, Cali y Barranquilla presentan patrones similares en las diferencias de estructura con Bogotá.

El presente documento está organizado en seis secciones, siendo esta *Introducción* la primera de ellas. La segunda *-Revisión de Literatura-* describe los principales hallazgos de trabajos previos sobre diferenciación regional de mercados laborales. La tercera sección corresponde al *Análisis Descriptivo* de las diferencias en ingresos laborales y características de los trabajadores entre Bogotá y las demás ciudades. La cuarta sección describe en detalle la *Metodología* utilizada para realizar la descomposición de las diferencias de ingresos laborales.

La quinta sección expone los *Resultados* obtenidos al realizar la descomposición en los dos casos mencionados; y por último, en la sexta sección se presentan las *Conclusiones*.

2. Revisión de Literatura

2.1. Teoría

A nivel internacional, en la literatura se distinguen dos corrientes teóricas que indagan por la diferenciación de mercados laborales. Por un lado, la teoría neoclásica admite diferencias de largo plazo entre ciudades, atribuibles a diferencias por características de la fuerza laboral y/o de *amenities*. Y por otro lado, la teoría de segmentación afirma que, existen diferencias entre mercados que van más allá de las admitidas por la teoría neoclásica, y que estas corresponden a diferencias en estructuras de los mercados. No obstante, aún no existe una teoría que logre explicar satisfactoriamente las diferencias de estructura, y desde el punto de vista del análisis empírico no se ha establecido un método estadístico para aislar las diferencias que admite la teoría neoclásica de las que admite la teoría de segmentación.

2.1.1. Teoría Neoclásica

Uno de los trabajos más recientes y relevantes en la literatura de mercados laborales locales es el de Moretti (2013), en donde se destacan las enormes diferencias entre las áreas metropolitanas de Estados Unidos en variables como los ingresos de los trabajadores, las tasas de innovación y productividad, y se realiza un análisis descriptivo sobre la convergencia de las variables hacia una media, a partir de la comparación de los niveles de estas en los años 1980 y 2000; además, se desarrolla un modelo teórico de equilibrio general, en el que las diferencias entre mercados locales se deben a desajustes por fuera del equilibrio y las diferencias en las *amenities* que ofrece cada área metropolitana.

El modelo de equilibrio general en Moretti (2013) va en línea con la mayoría de la literatura internacional y está basado en la visión neoclásica del mercado laboral según la cual existen diferenciales salariales compensatorios; no obstante, este modelo no tiene en cuenta la heterogeneidad de la oferta laboral, como determinante de los ingresos laborales, y poco se ajusta a la persistencia de los diferenciales observados en los datos. En oposición a esta

visión puramente neoclásica del mercado laboral, se encuentran artículos teóricos como el de Brueckner et al. (2000) en el que las diferencias de ingresos laborales se explican por la heterogeneidad de trabajadores, la cual es producto de las diferencias en habilidades y en localización.

2.1.2. Teoría de segmentación

Las diferencias entre mercados laborales se han estudiado principalmente desde la teoría de segmentación, y los mayores avances teóricos y empíricos de esta literatura han estado basados en la versión de un mercado dual de la teoría de segmentación, en donde se reconoce la existencia de dos sectores: un sector primario de buenos trabajos (sector formal) y otro de malos trabajos (sector informal).

Una completa revisión literaria sobre los estudios realizados desde los puntos de vista teórico y empírico acerca de segmentación de mercados laborales es el artículo de Taubman and Wachter (1987), quienes afirman que la teoría de segmentación nace debido a la carencia de explicaciones satisfactorias para las diferencias de ingresos laborales entre mercados laborales; pues si bien la teoría neoclásica admite una diferenciación de mercados a nivel geográfico, étnico, etc. la teoría de segmentación establece que los mecanismos de determinación de ingresos laborales son propios de cada segmento del mercado y de ahí surgen diferencias salariales que parecen inexplicables -desde la teoría neoclásica.

Debido a la atención que concentran las diferencias de mercados laborales en el ámbito internacional, los estudios empíricos en segmentación son numerosos pero aún no existe un consenso acerca de cuál es la metodología adecuada para probar que un mercado está segmentado, pues estadísticamente es difícil aislar las diferenciaciones que admite la hipótesis de segmentación y las que admite la teoría neoclásica. Por ejemplo, en el caso de mercados locales segmentados, no es claro cómo establecer si las diferencias salariales son producto de características particulares de cada unidad geográfica (diferenciación) o si son producto de las diferencias en los mecanismos y estructuras de fijación de ingresos laborales en cada unidad (segmentación).

2.2. Estudios para Colombia

Dada la falta de consenso a nivel teórico, se consultaron los estudios para Colombia sobre diferencias entre mercados laborales locales; encontrando que la mayoría se han enfocado en ingresos laborales y dentro de los métodos utilizados para contrastar empíricamente la hipótesis de diferenciación se identifican dos corrientes de análisis: la primera analiza el comportamiento de las series de tiempo, mientras que la segunda analiza las diferencias de ingresos laborales en niveles para determinado año. Además de estos estudios de ingresos laborales, se encuentran dos estudios que analizan la diferencia en tasas de desempleo entre ciudades: Díaz (2011); Cárdenas et al. (2014) y Díaz (2011), y uno que describe el comportamiento de las series de tiempo de algunas variables de mercado laboral por ciudades (Arango, 2011).

En línea con lo descrito, se presentan a continuación los estudios realizados sobre diferenciación en términos de ingresos laborales.

Series de tiempo e hipótesis de convergencia de ingresos laborales. El conjunto de estudios que adoptaron el enfoque de series de tiempo encontraron evidencia de tendencias comunes entre los ingresos laborales de las principales ciudades –al menos para trabajadores no calificados; no obstante, las diferencias en los niveles de ingresos laborales no tienden a desaparecer entre años. De esta manera, el análisis de las diferencias en niveles es relevante y se precisa que, bajo el enfoque de series de tiempo no se puede determinar si la causa de la diferenciación está en las características de los individuos, pues el uso de micro-datos no va más allá de la división de las series por calificación de la mano de obra.

De los estudios sobre ingresos laborales, los trabajos que se enmarcan dentro de esta corriente establecen que, hay integración de mercados cuando la diferencia de las series de ingresos laborales tiende a un diferencial de equilibrio, y para probarlo aplican técnicas de cointegración y pruebas de convergencia sobre las series de ingresos laborales; concluyendo que existe integración de mercados laborales, a continuación se describen los trabajos realizados en esta línea.

En un primer intento por estudiar el mercado laboral colombiano desde la perspectiva regional, Nupia (1997) utiliza las series de ingreso laboral diario de los trabajadores no calificados

en las 4 principales ciudades (Bogotá, Medellín y Cali) para el período 1976-1995 y a partir del análisis de cointegración concluye que hay integración entre los mercados laborales de Bogotá, Medellín y Cali, mientras que el de Barranquilla se encuentra totalmente segmentado.

Basados en este primer trabajo, Jaramillo et al. (2000) amplían el horizonte de análisis e incluyen además de pruebas de cointegración, pruebas de convergencia; ellos unifican las series de ingresos laborales para trabajadores no calificados de las 4 principales ciudades, entre 1945 y 1998, y concluyen que si bien Bogotá, Medellín y Cali presentan tendencias comunes, las únicas series que evidencian convergencia son las de Medellín y Cali.

Posteriormente, Galvis (2002) extiende el análisis a otras ciudades y para la fuerza laboral en su conjunto, no se limita a trabajadores no-calificados como los dos estudios anteriores. Él analiza las series de ingresos laborales para las 7 principales ciudades/AM (Bogotá, Medellín, Cali, Barranquilla, Manizales, Bucaramanga y Pasto) entre 1984 y 2000, y encuentra evidencia de convergencia regional para el mercado en su conjunto; al dividir las series por niveles educativos solo encuentra evidencia de convergencia para los niveles educativos más bajos.

Corte transversal y diferencias en los niveles de ingresos laborales. Todos los trabajos que han explorado las diferencias de ingresos laborales en niveles evidenciaron que estas no son puramente explicadas por las diferencias en características de la fuerza laboral. Este trabajo, a diferencia de los mencionados en esta sección, cuantifica las diferencias de ingresos más allá de las diferencias en la media, amplía el horizonte de análisis; y aunque al igual que en Quiñonez (2011) propone el uso de una descomposición del tipo Oaxaca-Blinder, desde el punto de vista del problema de identificación se sugiere el uso de otra metodología.

Romero (2008) usa los ingresos laborales reales de los ocupados entre enero de 2001 y junio de 2004 en 12 de las 13 ciudades/AM (no incluye Ibagué porque esta ciudad no cuenta con un índice de precios para deflactar los ingresos laborales en el período analizado) , estima ecuaciones mincerianas con la corrección del sesgo de selección de Heckman y emplea variables binarias por ciudad, tomando como referencia a Bogotá; él concluye que un gran porcentaje de las diferencias se explican por características de los individuos -excepto en el caso de

Barranquilla, e identifica tres clubes de ciudades por niveles de ingreso: 1) Bogotá, Medellín y Cali; 2) Bucaramanga, Villavicencio, Cartagena, Pereira y Cúcuta; y 3) Manizales, Pasto, Montería y Barranquilla. Por lo cual es este el primer trabajo en notar que las diferencias de ingresos laborales no se explican por diferencias en características de los trabajadores.

Resaltando la importancia de preguntarse por diferencias salariales más allá de la media, García et al. (2008) observan las series de ingresos laborales de los ocupados en las 7 principales ciudades en el período 2000-2005 y utilizan diversas pruebas estadísticas para establecer si las distribuciones de ingresos laborales son distintas entre sí, comparan año a año entre ciudades, entre sectores económicos y entre ciudad-sector económico, llegando a la conclusión de que las distribuciones de ingresos laborales difieren entre sí, aún después de hacer la corrección del sesgo de selección y de controlar por características como género, edad y educación. Así García et al. (2008) evidencian que las diferencias no se encuentran únicamente en la media de los ingresos laborales.

Posteriormente, motivados por la amplia documentación de la segmentación laboral del mercado colombiano, en términos de un sector formal e informal, Badillo et al. (2009) exploran la hipótesis de segmentación en dos dimensiones: intra e interregionalmente. Los autores utilizan las series de ingresos laborales de las 13 ciudades/AM para el período 2001-2006, estiman un modelo de ingresos laborales basado en ecuaciones mincerianas que incluye variables binarias por tipo de sector: informal (segmentación intrarregional), y variables binarias por ciudad (segmentación interregional), tomando como referencia a Bogotá; y a partir de la significancia de estas variables concluyen que en Colombia existe segmentación en ambas dimensiones, sus resultados se mantienen incluso después de utilizar el método de corrección del sesgo de selección de Heckman.

Por último, Quiñonez (2011) con los datos de ingresos laborales del segundo trimestre de 2008 para las 13 ciudades/AM utiliza la metodología de descomposición de ingresos laborales propuesta por Juhn et al. (1993) para explorar las diferencias de ingresos laborales entre Bogotá y cada una de las demás ciudades, llegando a la conclusión de que el componente de las diferencias asociado a lo no observado es relevante en la mayoría de los casos.

Finalizando los estudios sobre diferencias de ingresos laborales se encuentra Galvis (2010), siendo este el único que realiza un estudio desde las dos corrientes y se pregunta por las

diferencias de las series y de los niveles de ingresos laborales. En primera instancia, estima ecuaciones mincerianas para los asalariados de las siete principales ciudades entre 1984-2009 y analiza la convergencia de series; posteriormente, hace una descomposición de las variaciones de ingresos laborales no explicadas (variaciones obtenidas después de controlar por características observables) utilizando el Índice de Theil para estimar los componentes de variación intra e interregionales, encontrando evidencia de que las principales variaciones de ingresos laborales se encuentran explicadas por las variaciones interregionales. Así, se concluye que los diferenciales entre ciudades son significativos y en algunos casos se amplían a lo largo del tiempo.

Fuera del análisis de ingresos laborales existen dos estudios que analizan las tasas de desempleo incluyen herramientas de econometría espacial para identificar interdependencias entre las diferentes tasas. Díaz (2011) analiza los cambios de la tasa de desempleo de municipios colombianos entre 1993 y 2005, encontrando que la dispersión de las tasas creció en este período y que existe una correlación positiva entre las tasas de desempleo de municipios vecinos; adicionalmente, mediante un ejercicio de descomposición observa que el 66 % del cambio en las tasas de desempleo entre estos años está explicado por cambios en la tasa de desempleo de municipios cercanos. El otro estudio, elaborado por Cárdenas et al. (2014), utiliza un modelo MFACT (Multi Factor Analysis for Multiple Contingency Tables) para descomponer la varianza de las tasas entre regiones, encontrando que la varianza de las tasas de desempleo está asociada en un 76 % a un índice de calidad de la fuerza laboral; e identifica seis *clusters* de ciudades: el primero conformado por Quibdó, Cartagena, Riohacha y Valledupar; el segundo por, Popayán, Pasto, Montería, Neiva, Villavicencio y Sincelejo; el tercero por Barranquilla, Montería y Cartagena; el cuarto por Pereira, Armenia, Manizales, Ibagué, Cúcuta y Cali; el quinto por Bogotá, Tunja y Bucaramanga; y el sexto por Medellín. Y de manera más reciente, Arango (2011), quien usando los datos de las encuestas de mercado laboral realizadas por el DANE: ECH-GEIH, para el período 2001-2011 con 13 ciudades y para el período 2006-2011 con 23 ciudades, calcula el rango de variación promedio de los indicadores TGO, TO, TD, TS y TSO, y señala que este es superior a los 13 puntos porcentuales; por lo que el autor concluye que el mercado laboral colombiano es heterogéneo.

3. Análisis Descriptivo

3.1. Datos

Los datos utilizados en el presente trabajo provienen de la Gran Encuesta Integrada de Hogares GEIH para el período 2009-2015 ¹; esta encuesta es realizada por el Departamento Administrativo Nacional de Estadística DANE, es utilizada para realizar el cálculo de los principales indicadores del mercado laboral en Colombia, y además contiene información de la oferta laboral desagregada a nivel de individuo.

Para este trabajo se usaron los datos del entorno urbano de la muestra de las 13 principales ciudades, excluyendo del análisis los datos del entorno rural para evitar introducir una fuente de diferenciación distinta a la que se analiza en el presente estudio².

Se tomó la clasificación de la población y definiciones establecidas por el DANE para identificar a los individuos ocupados y clasificarlos en asalariados y cuenta propia. Además se tomaron variables de características individuales de los trabajadores y se definió una variable binaria que les asigna el carácter de formales si están afiliados a salud y pensiones ³. Así se conformaron cuatro grupos de trabajadores ocupados: formales-asalariados, formales-cuenta propia, informales-asalariados e informales-cuenta propia ⁴.

Por último, se toma la variable de ingresos laborales que es calculada por el DANE y reportada dentro de la encuesta GEIH y se utiliza el IPC por ciudades para expresar esta variable a precios de diciembre de 2014.

¹El período de referencia se eligió según la disponibilidad de datos. Por un lado, las series de ingresos laborales de la encuesta fueron empalmadas hasta 2008, y se elige 2009 como primer año para evitar posibles diferencias entre años que sean producto de los cambios realizados por el empalme; y por otro lado se finaliza en 2015 porque es el último año disponible.

²El Glosario de términos de la GEIH, elaborado por el DANE, define área metropolitana como “área de influencia que incluye municipios circundantes, que con la ciudad conforman un solo tejido urbano no discontinuo y han sido reconocidos legalmente”.

³Las preguntas de la encuesta utilizadas para ello fueron:

- Para definir la afiliación en salud: ¿... está afiliado, es cotizante o es beneficiario de alguna entidad de seguridad social en salud? (Instituto de Seguros Sociales - ISS, Empresa Promotora de Salud - EPS o Administradora de Régimen Subsidiado - ARS), y
- Para definir la afiliación en pensiones: ¿Está... cotizando actualmente a un fondo de pensiones?

⁴Se excluyen de la muestra los trabajadores que pertenecen a dos grupos.

3.2. Hechos Estilizados del Mercado Laboral Colombiano

Esta sección presenta un análisis descriptivo de las diferencias de ingresos laborales entre Bogotá y el conglomerado de las 12 ciudades; y entre las cinco ciudades, Bogotá, Medellín, Cali, Barranquilla y Bucaramanga. Se presenta el análisis para el total de los ocupados, y para los ocupados divididos en cuatro grupos: 1) formales asalariados, 2) formales cuenta propia, 3) informales asalariados y 4) informales cuenta propia. Esta división del análisis se basa en trabajos previos que han documentado la segmentación laboral del mercado colombiano en un sector formal y un sector informal (Badillo et al., 2009); y el trabajo de (García et al., 2009) quienes resaltan la diferencia entre los determinantes de ingresos laborales para trabajadores asalariados y cuenta propia.

Dado que los ingresos laborales están asociados a características de los trabajadores, fundamentalmente a variables de capital humano como la educación y experiencia; también se presentan estadísticas descriptivas de los años de escolaridad y de la edad de los trabajadores, el uso de esta última variable responde a la ausencia de datos de experiencia en la encuesta y teniendo en cuenta que varios de los estudios para Colombia calculan una variable ‘proxy’ de experiencia como la resta entre la edad y los años de educación.

Del análisis conducido se evidencia que existen diferencias tanto en los ingresos laborales como en las características de los trabajadores entre ciudades, que estas diferencias no son lineales a lo largo de la distribución y que no es posible adjudicar las diferencias observadas en ingresos laborales a diferencias en las características porque ambos tipos de diferencias exhiben patrones diferentes, incluso al dividir la muestra por grupos de trabajadores. Por ello, la discusión presentada sirve como motivación para conducir un ejercicio de descomposición de los diferenciales de ingresos laborales que contribuya a explicar los determinantes de las diferencias de ingresos laborales entre ciudades.

3.2.1. Bogotá y las 12 ciudades

La Figura 1 ilustra las diferencias de ingresos laborales, educación y edad entre Bogotá y las 12 ciudades, para 2009 y 2015 del total de trabajadores. Cada una de las cajas representa las siguientes estadísticas para Bogotá (cajas azules) y el resto de ciudades (cajas rojas): el promedio, dado por la línea que se encuentra al interior de cada caja; el percentil 25, dado

por el límite inferior de las cajas; el percentil 75, dado por el límite superior de las cajas⁵.

Al observar los ingresos laborales se evidencia que existe una diferencia a favor de Bogotá a lo largo de toda la distribución y para ambos años; aunque esta es mayor en el percentil 75, seguida de la diferencia en el percentil 25. Lo anterior implica, de una parte, que al conducir análisis de diferencias de ingresos laborales no es suficiente con caracterizar las diferencias en la media, pues la magnitud de estas varía a lo largo de la distribución; y por otro lado, que si las mayores diferencias se observan en el percentil 75 entonces es para el grupo de trabajadores calificados donde la diferencia es mayor. Aún más, estos patrones de la diferencia por ingresos se mantienen incluso al dividir la muestra de trabajadores en los cuatro grupos.

Sin embargo se precisa que, al hacer esta división las magnitudes de la diferencia varían significativamente entre los distintos grupos de trabajadores; razón por la cual no solo desde la discusión presentada sobre trabajos previos sino también desde el análisis conducido en esta sección se justifica el análisis para los diversos grupos de trabajadores.

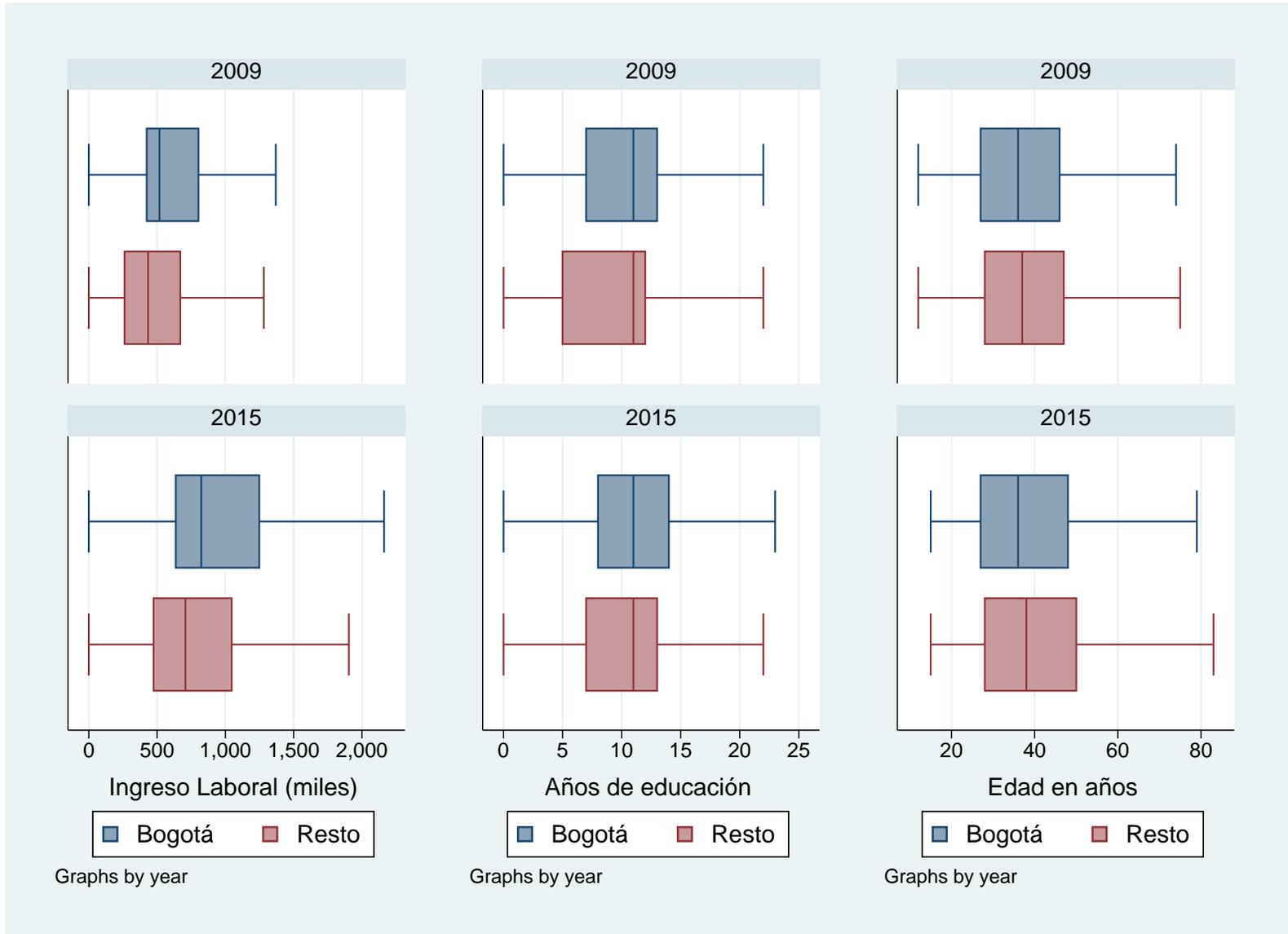
Acerca de las características de los trabajadores, se observa que en promedio los trabajadores en Bogotá cuentan con 0.5 años de educación más y son 0.9 años más jóvenes en relación con el grupo del resto de ciudades. Y aunque las diferencias en ingresos son distintas entre el percentil 25 y 75, tanto en años de educación como en edad la diferencia es exactamente igual en ambos percentiles.

Lo anterior pone en evidencia que las diferencias observadas en características no explican completamente las diferencias en ingresos laborales. Acerca de esto, también se destaca que al dividir la muestra para los diferentes grupos de trabajadores las diferencias en educación y edad son casi nulas para los grupos de trabajadores formales-asalariados e informales-asalariados y que es solo para el primero de estos grupos que se observa que la diferencia en ingresos también desaparece.

⁵La longitud de las líneas, está determinada por los límites superior e inferior que corresponden a valores adyacentes de los rangos inter-cuartílicos y son un indicador de la varianza entre cada uno de los grupos.

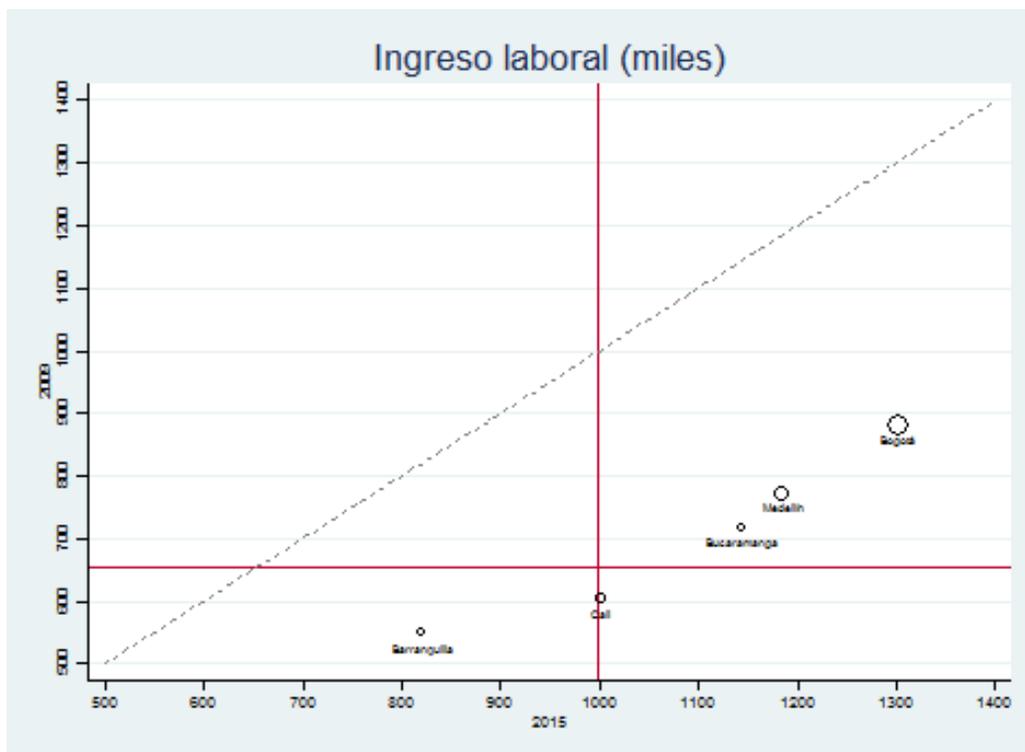
Figura 1: Ingresos laborales, Años de Educación y Edad - Bogotá y las 12 ciudades - Total de ocupados.

15



3.2.2. Bogotá y otras ciudades principales

Figura 2: Ingresos laborales - Bogotá y otras ciudades principales



Diferencias en la media. Al realizar la comparación entre Bogotá, Medellín, Cali, Barranquilla y Bucaramanga se muestra que hay un amplio rango de variación, como ya había sido notado por Arango (2011) y en adición se observa que, no hay evidencia de la convergencia de los ingresos laborales hacia una media y las posiciones relativas de las ciudades para la muestra del total de ocupados se mantienen; es decir, aquellas ciudades que tenían niveles por encima (debajo) de la media en 2009 tienden a tener niveles por encima (debajo) de la media en 2015.

También se observa persistencia en el tiempo de las diferencias en los niveles de ingresos laborales inclusive cuando se tienen en cuenta las dos dimensiones de diferenciación (formal-informal, y asalariado-cuenta propia); por lo que el análisis de esta sección junto con estudios previos que abordan la diferenciación de mercados laborales, motivan la estimación de los resultados para los diferentes grupos de trabajadores.

La Figura 2 reporta los niveles promedio de ingresos laborales para 2009 y 2015, líneas rojas horizontal y vertical, respectivamente; desagregando los niveles de ingresos laborales en cada

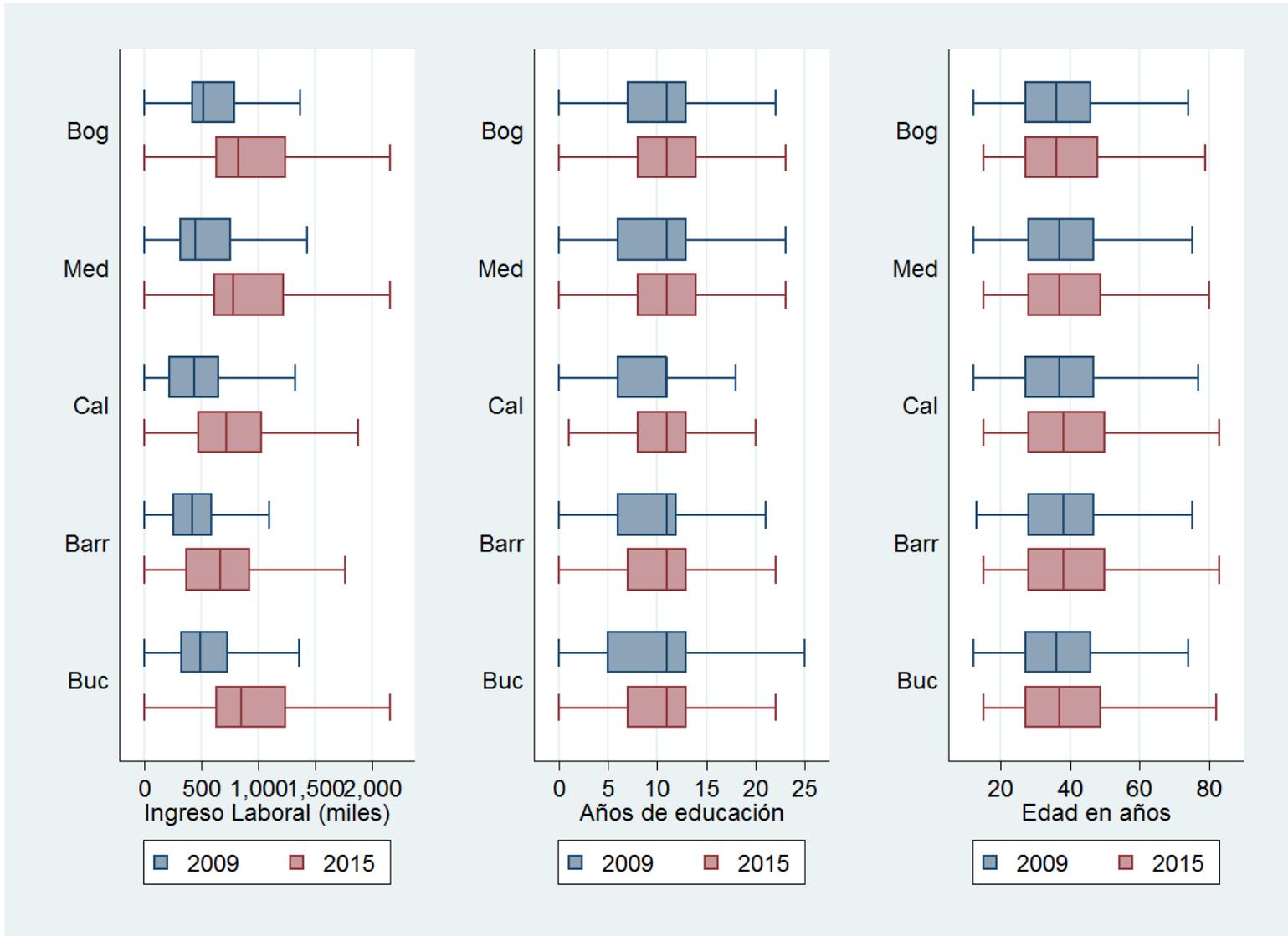
una de las ciudades. Además, se muestra el tamaño relativo de los mercados laborales de cada ciudad, representado por el tamaño de los círculos.

Al observar los niveles de ingresos laborales, se evidencia un rango significativo de variación entre las cinco ciudades para ambos años pues, en 2009 Bogotá con un una media de ingresos laborales de \$800.000 es la ciudad con mayores ingresos laborales, mientras que Barranquilla con \$550.000 es la ciudad con menores ingresos laborales. Y en 2015, ambas ciudades conservan sus posiciones relativas con ingresos de \$1.250.000 y \$820.000.

Para el período 2009-2015, se observa un aumento promedio del 35 %; esta magnitud de varación es similar para cuatro de las cinco todas las ciudades, siendo Barranquilla la ciudad que presenta una varación menor. Adicionalmente, se distinguen dos grupos de ciudades: Bogotá, Medellín y Bucaramanga con niveles de ingresos laborales mayores que el promedio en ambos años; y que Cali y Barranquilla, con niveles de ingresos laborales menores que el promedio.

Al dividir el análisis en los cuatro grupos de trabajadores, aunque se sigue evidenciando heterogeneidad de los ingresos laborales por ciudades. Se precisa que de los cuatro grupos de trabajadores, el que mayor semejanza presenta con el total de los ocupados es el grupo de trabajadores formales-asalariados, y esto es porque es el grupo que representa la mayoría de los ocupados (ver Anexo Figura 9); el único grupo en el que disminuyó el rango de variación entre años fue el de formales-cuenta-propia, y es este el grupo en donde las variaciones temporales de los ingresos son más heterogéneas, haciendo que las posiciones de las ciudades con respecto al promedio varíen entre años (ver Anexo Figura 10).

Figura 3: Ingresos laborales, Años de Educación y Edad - Bogotá y otras ciudades principales



Diferencias a lo largo de toda la distribución. En los percentiles 25 y 75, al igual que ocurre en el caso de la media, Medellín y Bucaramanga son las ciudades cuyos niveles de ingresos laborales están más cerca a los de Bogotá; no obstante, la magnitud de la diferencia no es lineal y difiere a lo largo de toda la distribución. Acerca de las diferencias en las características observadas de los trabajadores, se observa que la media de años de educación no es significativamente distinta entre ciudades y que esta no varía entre años. Y aunque los patrones de diferenciación son mayores en los percentiles 25 y 75, que en la media, estos no reflejan los patrones de diferenciación de los ingresos laborales (ver Figura 3).

Y al dividir la muestra de ocupados en los cuatro grupos, se observa que las diferencias entre ciudades y a lo largo de la distribución continúan siendo distintas. Aún más, las diferencias en características por grupos tampoco replican los patrones de diferenciación por ingresos (ver Anexo Figura 11 y Figura 12).

4. Metodología

En esta sección se describe la estrategia econométrica utilizada para realizar la descomposición del diferencial de ingresos laborales entre Bogotá y las demás ciudades. Inicialmente, se presenta una descripción general de los métodos de descomposición del tipo Oaxaca-Blinder y posteriormente se describe en detalle la metodología de *Reweighting Regression* (Fortin et al., 1996), que es la utilizada en este estudio para calcular los resultados de la descomposición presentados en la siguiente sección.

Los métodos de descomposición del tipo Oaxaca-Blinder. Partiendo de las distribuciones observadas $f_{(Y_g)}$ de los ingresos en las ciudades g , es posible calcular la diferencia de ingresos laborales entre Bogotá y otras ciudades, y evaluar estas diferencias a lo largo de toda la distribución.

En particular, los métodos de descomposición del tipo Oaxaca-Blinder se concentran en la diferencia en algún estadístico de la distribución,

$$\Delta_O^v = v(f_{Y_B|D_B}) - v(f_{Y_A|D_A})$$

Donde

- $f_{(Y_g|D_g)}$ son las distribuciones observadas de ingresos laborales en Bogotá ($g = B$) y otra ciudad ($g = A$).
- $v(f_{(Y_g|D_g)})$ es cualquier estadístico obtenido a partir de esas distribuciones.

Y se descompone la diferencia total Δ_O^v en: aquella parte que corresponde a diferencias en las características observadas de la fuerza laboral Δ_X^v (efecto composición) y aquella parte que corresponde a diferencias en los retornos a estas características, que corresponde a lo que se conoce en la literatura de métodos de descomposición como efecto estructura de ingresos laborales Δ_S^v , por lo que se tiene

$$\Delta_O^\mu = \Delta_X^\mu + \Delta_S^\mu$$

En el caso simple de la media ($v = \mu$), los términos de la descomposición de Oaxaca-Blinder estarían dados por:

$$\Delta_X^\mu = \mu(f_{Y_B|D_B}) - \mu(f_{Y_B^C:X=X|D_A})$$

$$\Delta_S^\mu = \mu(f_{Y_B^C:X=X|D_A}) - \mu(f_{Y_A|D_A})$$

Es decir que para calcular los términos asociados a la descomposición se requieren:

- $\mu(f_{Y_g|D_g})$ que corresponde a la media de los ingresos laborales en cada ciudad $g = A, B$, y se obtienen de las distribuciones observadas $f_{Y_A|D_A}$ y $f_{Y_B|D_B}$.
- $\mu(f_{Y_B^C:X=X|D_A})$ que sería la media de ingresos laborales de los trabajadores de la ciudad A , si sus características se remuneraran según la estructura de ingresos laborales de Bogotá ⁶. El cálculo de este término implica la estimación de una distribución contrafactual $f_{Y_B^C:X=X|D_A}$ que no es observada; y las distintas alternativas de estimarla dan

⁶Aquí se está asumiendo que el grupo de referencia es el grupo A . La descomposición también puede realizarse tomando como referencia el grupo B , en cuyo caso se estimaría $\mu(f_{Y_B^C:X=X|D_A})$ en lugar de $\mu(f_{Y_B^C:X=X|D_A})$

origen a diversas técnicas de descomposición, siendo la estimación de este término un ejercicio análogo a la estimación de un contra-factual en los ejercicios de evaluación de impacto.

Si se considera que las distribuciones observadas de ingresos laborales $f_{Y_g|D_g}(y)$ están dadas por las características de los trabajadores en cada ciudad $g = A, B$, aplicando la ley de probabilidades iteradas, se tiene que:

$$f_{Y_g|D_g}(y) = \int f_{Y_g|X,D_g}(y|X = x)df_{X|D_g}$$

Seguendo a Mincer (1974), si se tiene en cuenta que los ingresos laborales del individuo i dependen de variables observadas X_i y variables no observadas ε_i , mediante las funciones de estructura de ingresos laborales $m(\cdot)$ en cada ciudad g , es decir que, la ecuación que describe el ingreso laboral de un trabajador adopta la siguiente forma: $Y_{gi} = m_g(X_i, \varepsilon_i)$. Entonces es posible afirmar que:

$$f_{Y_g}(y) = \int Pr(m_g(X, \varepsilon) = y|X = x, D_g = 1)df_{X|D_g}$$

Y las diversas alternativas de manipular los dos términos de la ecuación para estimar la distribución contrafactual, dan origen a las distintas técnicas de descomposición.

Acerca del problema de elección de una técnica, el artículo de Fortin et al. (2011) expone las diversas técnicas de descomposición existentes y se presenta una completa discusión sobre qué método elegir desde una aproximación econométrica al problema de identificación; y el artículo de Goraus et al. (2013) realiza una evaluación empírica de las diversas técnicas disponibles, a partir de estos artículos se determina que la técnica adecuada para realizar el presente ejercicio es la de *Reweighting Regression* y se describe a continuación.

4.1. Método econométrico a utilizar: *Reweighting Regression*

La estimación de la distribución contra-factual bajo esta metodología propone, en lugar de integrar la distribución de ingresos laborales de Bogotá usando la distribución marginal de cada una de las variables en el vector de características X de las otras ciudades A , utilizar un factor de pesos, de tal forma que la distribución contra-factual estaría dada por la expresión:

$$f_{Y_B^C: X=X|D_A}(y) = \int f_{Y_B|X, D_B}(y|X=x)df_{X|D_A} = \int f_{Y_B|X, D_B}(y|X=x)\varphi(X)df_{X|D_B}$$

Donde:

$$\varphi(X) = df_{X|D_A}/df_{X|D_B}$$

Y utilizando la regla de Bayes:

$$Pr(X|D_B = 1) = \frac{Pr(D_B = 1|X)df(X)}{\int Pr(D_B = 1|X)df(X)} = \frac{Pr(D_B = 1|X)}{Pr(D_B = 1)}$$

Análogamente:

$$Pr(X|D_B = 0) = \frac{Pr(D_B = 0|X)df(X)}{\int Pr(D_B = 0|X)df(X)} = \frac{Pr(D_B = 0|X)}{Pr(D_B = 0)}$$

De esta forma:

$$\begin{aligned} \varphi(X) &= \frac{Pr(X|D_B = 0)}{Pr(X|D_B = 1)} \\ &= \frac{Pr(D_B = 0|X)/Pr(D_B = 0)}{Pr(D_B = 1|X)/Pr(D_B = 1)} \end{aligned}$$

Por lo tanto, para obtener $\varphi(X)$ se estima un modelo que calcule $P(D_B = 1|X)$ -logit o probit por ejemplo; y se obtiene directamente de la muestra el valor de $P(D_B = 1)$. Los valores de $P(D_B = 0|X)$ y de $P(D_B = 0)$ se obtienen simplemente restando las anteriores probabilidades a 1, en cada caso ⁷.

Desde esta aproximación, la estimación del contra-factual requiere que se cumpla el supuesto de soporte común, de manera que se garantice que $P(D_g = 1|X) \neq 0$. Bajo este supuesto, con la estimación del contra-factual es posible calcular todos los términos de la diferencia, para algún estadístico de interés v , de tal manera que

$$\Delta_O^{v(f_Y)} = \Delta_X^{v(f_Y)} + \Delta_S^{v(f_Y)}$$

⁷ $P(D_B = 0|X) = 1 - P(D_B = 1|X)$ y $P(D_B = 0) = 1 - P(D_B = 1)$

El término $\Delta_X^{v(f_Y)} = v(f_{Y_B}) - v(f_{Y_B^C: X=X|D_A})$ refleja las diferencias en las características, pues corresponde a la diferencia de ingresos entre los trabajadores de Bogotá y los de otras ciudades A , si todos se localizaran en Bogotá. Es importante notar que esto se logra sin tener ningún supuesto adicional sobre la estructura de los datos. Aún más, bajo el supuesto de ignorabilidad, el cual se cumple cuando las distribuciones condicionales de las características no observadas no son estadísticamente diferentes entre las ciudades para las que se hace la comparación, este término refleja únicamente las diferencias en las características observadas. Mientras que $\Delta_S^{v(f_Y)} = v(f_{Y_B^C: X=X|D_A}) - v(f_{Y_A})$ se interpreta como el efecto estructura de ingresos laborales y refleja las diferencias en los retornos a las características observadas, pues corresponde a la diferencia de ingresos laborales que tendrían los trabajadores de otras ciudades A si ellos localizaran en Bogotá y no en la ciudad actual. De forma general, la diferencia en los retornos a las características observadas está determinada por las diferencias en las funciones de estructura de ingresos laborales m_g .

El problema de identificación. Con el fin de tener resultados robustos se adoptan en el presente ejercicio las siguientes consideraciones:

- Dividir la muestra para comparar grupos de trabajadores homogéneos y minimizar los errores que son producto del sesgo de selección, esta división es motivada por los estudios previos y el análisis descriptivo presentado en la sección 2.2. Con esto se busca garantizar que se cumpla el supuesto de ignorabilidad.⁸
- Explorar las diferencias más allá de la media, por lo que la descripción de los resultados obtenidos estará en función de los diferenciales a lo largo de la distribución.
- Elegir una especificación econométrica que más se ajuste a los datos; aunque el método no asume una forma funcional predeterminada de la relación entre la variable de ingresos laborales y las variables independientes, si se debe tener en cuenta que las variables

⁸Aunque también existe la posibilidad de estimar una primera etapa y realizar una corrección por sesgo de selección, para este caso, dadas las diferencias estructurales entre los diversos grupos de trabajadores, se considera una mejor alternativa dividir la muestra. Esto porque las diferencias entre grupos de trabajadores obligan a realizar un método de corrección de selección multinomial, y aún no existen estudios que validen económicamente los resultados de la descomposición bajo este tipo de correcciones. Aún más, en caso de realizar una corrección de selección bivariada existe un problema a la hora de elegir la variable instrumental porque esta puede estar corrigiendo un tipo de sesgo, pero agregando otro.

independientes incluidas son aquellas que mejor se ajustan a los datos observados. Para esto se tomó como referencia el trabajo de Londoño (2013), quien a partir del uso del método de promedios bayesianos con los datos de la GEIH estima ecuaciones mincerianas y concluye que las variables con mayor probabilidad de ser incluidas en el modelo son años de educación, género, edad y estado civil.

Por último, a continuación se exponen las ventajas y debilidades de esta técnica.

Ventajas. Las principales ventajas son su simplicidad de implementación y eficiencia (Busso et al., 2014; Firpo, 2007). En relación con otras técnicas de descomposición, es la que menos supuestos impone sobre los datos y es una de las más informativas; pues no impone una forma funcional de la ecuación de ingresos laborales en términos de las variables observadas de la fuerza laboral y permite evaluar las diferencias a lo largo de toda la distribución, por lo cual se realiza la descomposición para la media, el percentil25 y el percentil75.

Al compararla con las técnicas que realizan estimaciones de ecuaciones mincerianas y se concentran en el cálculo de los estadísticos de las distribuciones, mirar lo que ocurre a lo largo de toda la distribución de ingresos laborales es una mejor alternativa que concentrarse únicamente en algún estadístico porque cualquier estadístico se encuentra influenciado por el efecto del salario mínimo.

Aún más, las técnicas que utilizan regresiones lineales utilizan la ley iterada de valores esperados y se basan en el supuesto de la media no condicional de los parámetros β :

$$E(Y) = E_X[E(Y|X)] = E(X)\beta$$

Aquí, β refleja el incremento en el valor medio de X sobre la media no condicional de Y . Por lo tanto, se asume un efecto lineal a lo largo de toda la distribución y cualquier error de especificación que afecte el valor de β , también afecta los resultados de la descomposición.

Aunque el problema del efecto lineal se puede solucionar utilizando técnicas que utilicen la regresión por cuantiles, en estos casos se deben realizar varios supuestos para obtener los dos términos de la descomposición porque la ley de iterada de valores esperados no aplica así que:

$$Q_\tau \neq E_X[Q_\tau(X)] = E(X)\beta_\tau$$

Si bien existen técnicas como la de Machado and Mata (2005) y la de Fortin et al. (2009) que solucionan este inconveniente, es necesario realizar supuestos adicionales sobre la estructura de los datos; inclusive en el mejor de los escenarios, bajo el cual todos los supuestos se cumplen para llevar a cabo una técnica que involucre la estimación de ecuaciones mincerianas, la descomposición para estadísticos diferentes a la media involucra términos de interacción para los cuales no hay interpretación.

Debilidades. En términos de implementación se encuentran los siguientes inconvenientes: 1) la dificultad en el análisis porque sólo permite comparar por pares de ciudades (grupos de ciudades) y 2) no se han desarrollado técnicas estándar para establecer la significancia estadística de cada uno de los componentes del diferencial. A continuación se explica cómo se propone abordar estas limitaciones en el presente estudio.

Sobre la comparación por pares de ciudades, no parece haber una solución estadística y claramente esto limita los análisis de diferenciación a un número de ciudades. Como solución a esta limitación del método elegido, se propone comparar a Bogotá, que es la ciudad principal (mayores ingresos laborales y mayor número de personas que participan en el mercado laboral), con el resto de ciudades y con las otras 4 ciudades con mayor número personas activas.

Y sobre la significancia de los diferenciales, éste es un tema de desarrollo investigativo reciente por lo que aún no existe una técnica universalmente aceptada; no obstante, debido a que el uso de la técnica de *Reweighting Regression* implica la estimación de las distribuciones de ingresos laborales en cada ciudad y la distribución contra-factual, se propone aplicar el test de Kolmogorov-Smirnov para determinar si las distribuciones de ingresos laborales y su contra-factual son estadísticamente diferentes o no. Una vez demostrada la significancia estadística de las diferencias, se cuantifica la magnitud de la diferencia y sus componentes en diferentes puntos de la distribución de ingresos.

4.2. Ecuaciones a estimar

En línea con lo planteado en la sección anterior se estima un modelo en dos etapas.

Primera Etapa Se calculan los factores de ponderación, que son requeridos para construir el contra-factual, mediante la estimación de un modelo probit de la siguiente forma:

$$Ciudad_i = edad_i + educa_i + gen_i + horas_i$$

Donde

- $Ciudad_i$ es una variable categórica que toma el valor de 1 si el individuo observado se localiza en la ciudad de referencia Bogotá B .
- $edad_i$ es la edad en años del individuo.
- $educa_i$ es una variable continua que indica los años de educación del individuo.
- gen_i es una variable binaria que toma el valor de 1 si el individuo es hombre.
- $horas_i$ es una variable continua que indica el número de horas trabajadas.

Segunda Etapa Se estima la distribución de ingresos laborales empírica para cada una de las ciudades, y suponiendo en cada ciudad $g = A, B$ la ecuación que determina el ingreso laboral de un individuo y_{gi} tiene la siguiente forma:

$$y_{gi} = m_g(X_i, \varepsilon_i), i \in g$$

Se construye la distribución hipotética de los ingresos laborales que tendrían los trabajadores de A si estos se localizaran en Bogotá, es decir que la ecuación que determina el ingreso laboral de los individuos en esta ciudad hipotética tiene la siguiente forma:

$$y_{gi}^C = m_B(X_i, \varepsilon_i), i \in A$$

Y posteriormente se estiman los términos de la descomposición.

5. Resultados

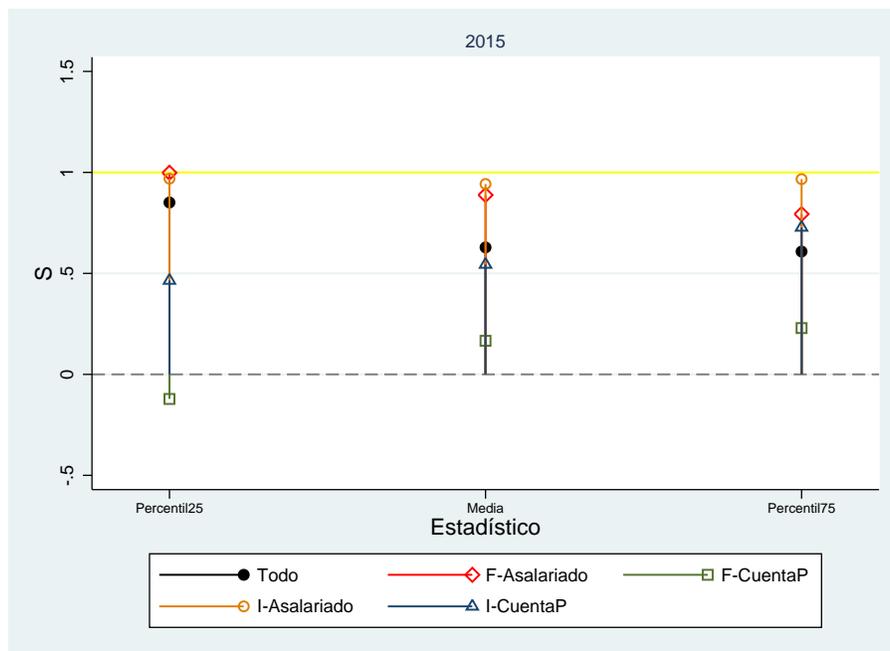
En esta sección se presentan los resultados obtenidos al conducir el ejercicio de descomposición para la media y los percentiles 25 y 75.

Y con el fin de determinar la importancia de la diferencia en estructura $\Delta_S^{v(f_Y)}$ para explicar la diferencia total $\Delta_O^{v(f_Y)}$, recordando que, según la discusión presentada en este trabajo, las diferencias en estructuras de ingresos corresponden a las diferencias en los retornos a las características observadas de la fuerza laboral, se calcula S como la proporción que representa la diferencia en estructura de la diferencia total, es decir:

$$S = \frac{\Delta_S^{v(f_Y)}}{\Delta_O^{v(f_Y)}}$$

5.1. Bogotá y las otras 12 ciudades

Figura 4: Diferencia en estructura en 2015 - Bogotá y las 12 ciudades.



En esta sección se presentan los resultados obtenidos al descomponer la diferencia de ingresos laborales entre Bogotá y las otras 12 ciudades, siendo Bogotá la ciudad de referencia. Se encuentra que las diferencias en estructura están a favor Bogotá y explican más del 50 % de la diferencia total en la mayoría de los casos.

5.1.1. Diferencias en 2015

Según los resultados de la Figura 4, en la media se observa que las diferencias en estructura están a favor de Bogotá ($S > 0$, dado que $\Delta_O^{v(f_Y)} > 0$) y explican el 56 % de la diferencia

observada entre los ingresos laborales de Bogotá y el resto de ciudades. No obstante, este porcentaje varía significativamente entre grupos de trabajadores; identificando que las mayores diferencias en estructura son para los trabajadores asalariados (valores más altos de S), y únicamente para el grupo de trabajadores formales-cuenta propia, que es grupo más pequeño (ver Anexo Figura 7), la diferencia en estructura corresponde a menos de la mitad de la diferencia total ($S < 0,5$).

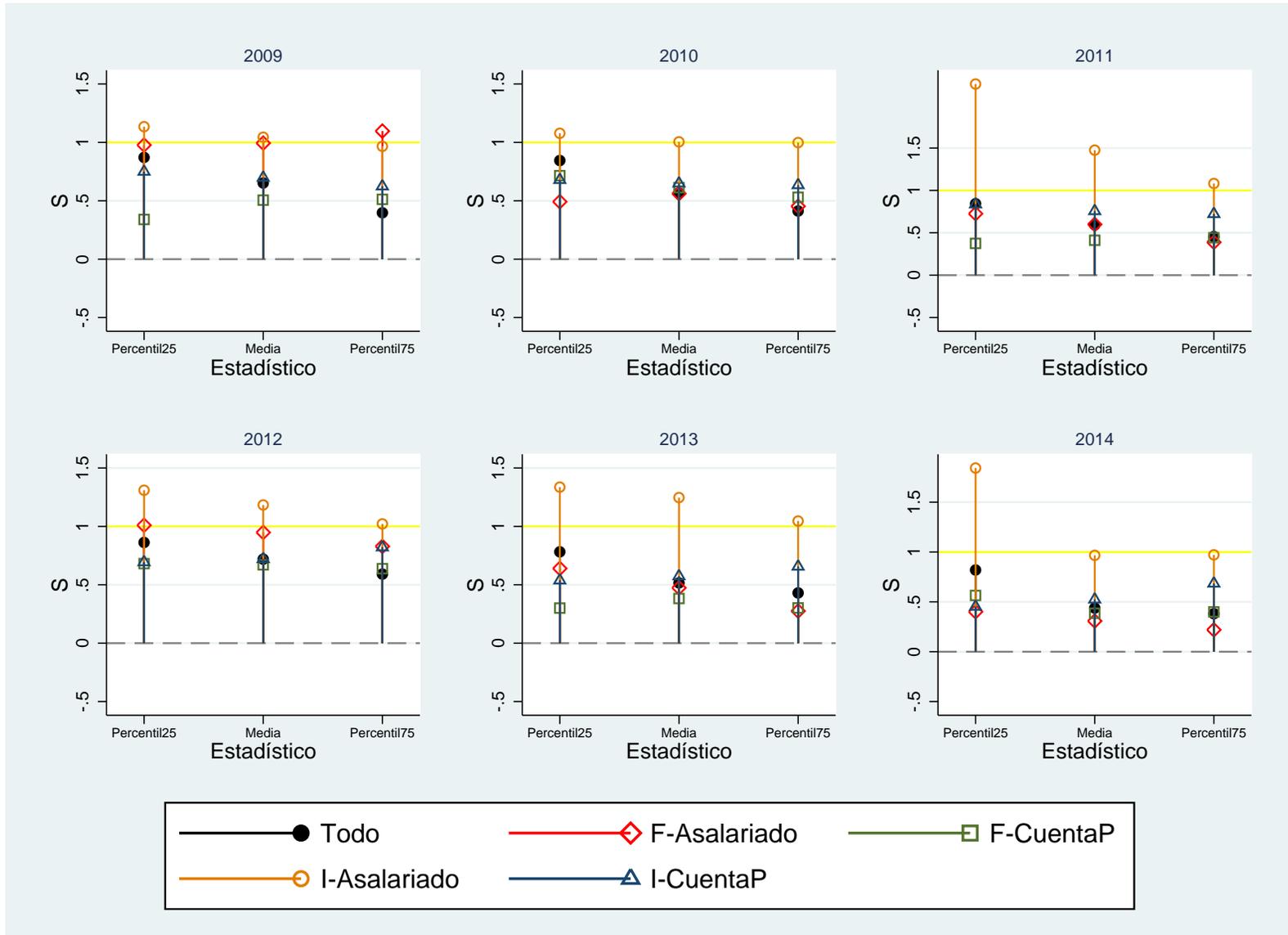
Y acerca de las diferencias a lo largo de la distribución, también en los percentiles 25 y 75 los trabajadores asalariados tienen mayores valores de S . Adicionalmente, por grupos se establece una asociación positiva entre S y los percentiles para los trabajadores cuenta propia (tanto formales como informales), indicando que S aumenta conforme aumentan los ingresos laborales; para los trabajadores formales-asalariados la relación entre el percentil y S entre los formales es negativa; y por último S no varía significativamente a lo largo de la distribución para los trabajadores informales-asalariados.

5.1.2. Diferencias en el tiempo

A través de los años, los mayores valores de S están en el percentil 25, es decir que, para ingresos laborales menores la diferencia en estructura representa una mayor proporción de la diferencia total; y entre grupos, los mayores valores de S son para el grupo de trabajadores informales-asalariados (ver Figura 5).

Y al comparar los valores de S entre 2009 y 2015, para los grupos de trabajadores formal-asalariado e informal-cuenta propia, que son los grupos más grandes, y el total de ocupados se evidencia una disminución de la diferencia en estructura, especialmente en el percentil 25 y la media; aunque está sigue explicando más de la mitad de la diferencia, como se afirmó anteriormente.

Figura 5: Diferencia en estructura en el tiempo - Bogotá y las 12 ciudades.



5.2. Bogotá y otras ciudades principales

En esta sección se presentan los resultados obtenidos al realizar la descomposición entre Bogotá y cada una de las otras 4 ciudades, para el total de la muestra y la muestra dividida en cuatro grupos.

5.2.1. Diferencias en 2015

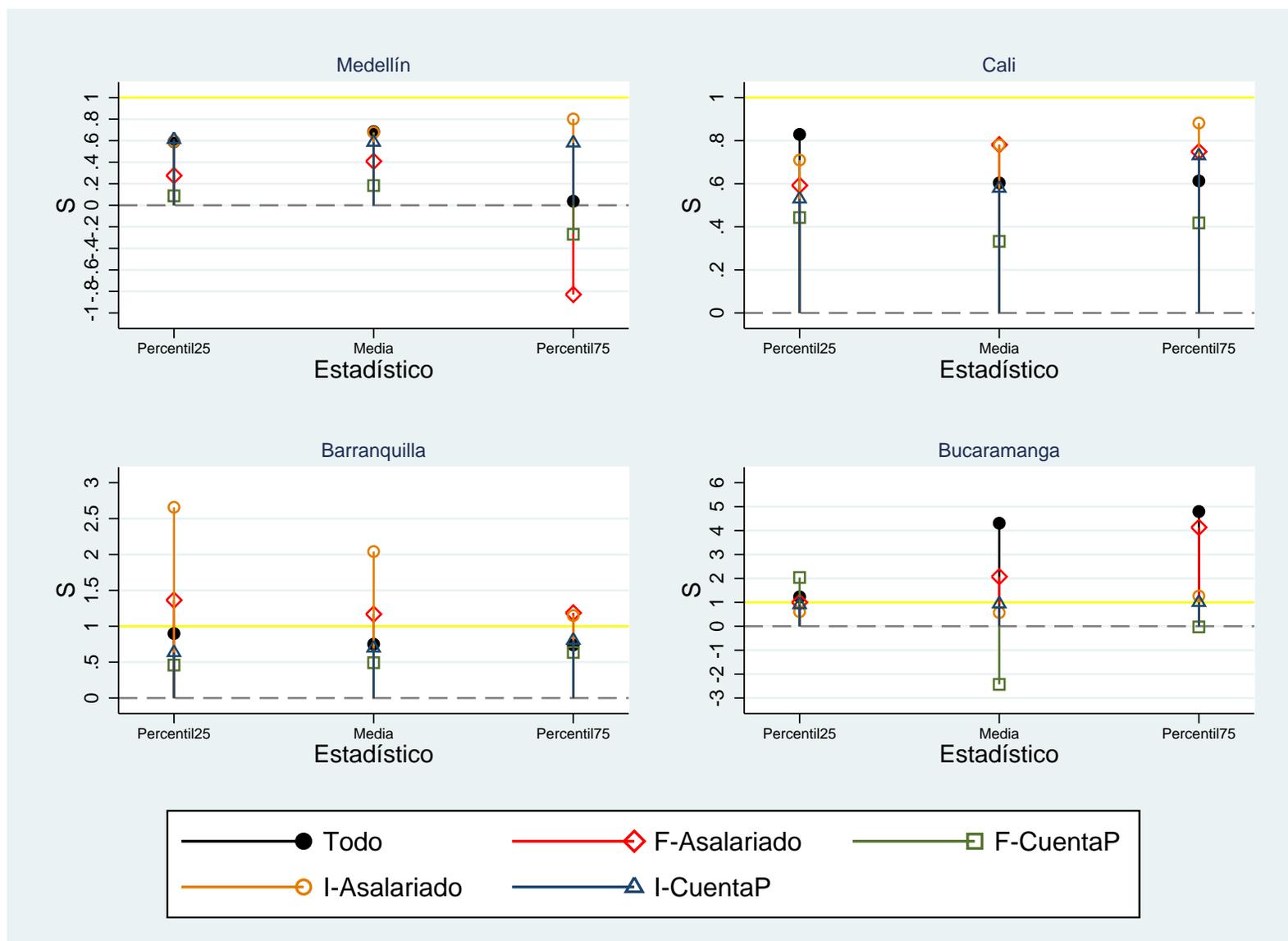
Medellín. Las diferencias en estructura están a favor de Bogotá en el percentil 25 y la media; mientras que en percentil 75 para los trabajadores formales están a favor de Medellín (Figura 6). Esto indica que las características de los trabajadores formales de mayores ingresos son mejor remuneradas en Medellín aunque los de Bogotá están mejor calificados, haciendo que la diferencia total sea a favor de la ciudad capital ($S < 0$ y $\Delta_O^{v(f_Y)} > 0$). Y entre grupos, se observa que las mayores diferencias están para los trabajadores formales.

Cali. Con Cali las diferencias están a favor de Bogotá a lo largo de toda la distribución ($S > 0$ y $\Delta_O^{v(f_Y)} > 0$), siendo ligeramente mayores en el percentil 75 (Figura 6). Y entre grupos, las menores diferencias en estructura son para trabajadores cuenta propia.

Barranquilla. De manera similar a lo que ocurre con Cali, las diferencias en estructura están a favor de Bogotá y son menores para los grupos de trabajadores cuenta propia (Figura 6). Sin embargo, la magnitud de S es significativamente mayor al comparar Barranquilla; e incluso, para los grupos de trabajadores asalariados $S > 1$.

Bucaramanga. De las ciudades analizadas, es para la cual la diferencia en estructura representa una mayor proporción de la diferencia total (Figura 6). En todos los casos la diferencia observada está a favor de Bogotá, pues aunque para los trabajadores formales-cuenta propia $S < 0$ en la media, se tiene que la diferencia total en la media también es negativa ($\Delta_O^{\mu(f_Y)} < 0$).

Figura 6: Diferencia en estructura (2015) - Bogotá y otras ciudades principales.



5.2.2. Diferencias en el tiempo.

Acerca de los patrones de diferenciación entre 2009-2014 (ver Anexo Figura 13, Figura 14, Figura 15, Figura 16, Figura 17 y Figura 18) se evidencia que:

Medellín. En relación con las otras ciudades, es la que menores diferencias tiene con Bogotá. Y la diferencia en estructura a favor de Medellín para los trabajadores formales de ingresos en el percentil 75, empieza a ser a favor de esta ciudad en 2013, año a partir de cual se ha ampliado.

Cali. El grupo de trabajadores informales-asalariados es para el cual se observa mayor variabilidad entre años; sin embargo, esta variabilidad no afecta de forma significativa las diferencias para el total de la muestra porque este es uno de los grupos de trabajadores más pequeños.

Barranquilla. Los mayores valores de S son para los trabajadores informales-cuenta; aunque estos no se reflejan en los valores de S para el total de los ocupados por el tamaño relativo de este grupo.

Bucaramanga. En relación con las otras ciudades, es la que mayores diferencias tiene con Bogotá. Los valores observados de S en el percentil 75 en 2011 y 2012, que representan cambios abruptos con respecto a sus niveles en otros años, se explican por a una reducción significativa de la diferencia total de los ingresos laborales con Bogotá para este percentil.

6. Conclusiones

Estudios previos para Colombia han documentado la persistencia en el tiempo de las diferencias regionales de ingreso en Colombia, y la literatura internacional ha encontrado evidencia de los efectos negativos que tienen las disparidades regionales sobre el crecimiento económico y el bienestar agregado. Dado lo anterior, y teniendo en cuenta que un alto porcentaje de los ingresos totales de las personas provienen de los ingresos laborales, se hace necesario el diseño

de instrumentos de política pública que sean efectivos a la hora de reducir las disparidades de ingresos laborales.

No obstante, previo al diseño de instrumentos efectivos es fundamental conocer los determinantes de estas diferencias; y para ello el presente trabajo propone el uso de una metodología de descomposición del tipo Oaxaca-Blinder la cual permite descomponer las diferencias de ingresos laborales en la parte que es atribuida a diferencias en las características de la fuerza laboral y la que es atribuida a diferencias en los retornos a esas características.

Al conducir el ejercicio de descomposición propuesto para Colombia, se logró determinar que las diferencias observadas en ingresos laborales de las demás ciudades con Bogotá se explican principalmente por la forma en cómo se remuneran las características de los trabajadores según la estructura de ingresos en cada ciudad. En línea con lo expuesto, los resultados obtenidos indican que si se quieren formular instrumentos de política pública más eficientes a la hora de reducir las disparidades de ingresos entre ciudades no es suficiente con elevar los años de educación de los trabajadores; y aún más, indican que es necesario indagar sobre las razones por las cuales las estructuras de ingresos entre ciudades difieren.

Adicionalmente se establece que, es posible caracterizar las diferencias entre los ingresos laborales de las ciudades haciendo uso de los métodos de descomposición del tipo Oaxaca-Blinder y desde una aproximación al problema de identificación, la cual tiene una analogía con el problema de identificación de los métodos de evaluación de impacto. Por ejemplo, si se requieren establecer grupos de análisis entre ciudades en base a las diferencias en estructura, es posible conducir ejercicios análogos al presentado con diferentes enfoques en los varíen las ciudades o grupos de ciudades a ser comparados.

Referencias

- Arango, L. (2011). Mercado de trabajo de Colombia: suma de partes heterogéneas. Borradores de Economía 671, Banco de la República de Colombia.
- Badillo, E., Ortiz, C., and Uribe, J. (2009). Segmentación inter e intrarregional en el mercado laboral urbano de Colombia, 2001-2006. *Ensayos sobre Política Económica*, 27(58):194–231.
- Bérnabou, R. (1996). Equity and Efficiency in Human Capital Investment: The Local Connection. *Review of Economic Studies*, 63(2):237–264.
- Blinder, A. (1974). On Dogmatism in Human Capital Theory. Working Papers 434, Princeton University, Department of Economics, Industrial Relations Section.
- Bonilla, L. (2008). Diferencias regionales en la distribución del ingreso en Colombia. Documentos de trabajo sobre Economía Regional 005284, Banco de la República de Colombia.
- Brueckner, J., Thisse, J.-F., and Zenou, Y. (2000). Local Labour Markets, Job Matching and Urban Location. CEPR Discussion Papers 2612, C.E.P.R. Discussion Papers.
- Busso, M., DiNardo, J., and McCrary, J. (2014). New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators. *The Review of Economics and Statistics*, 96(5):885–897.
- Cárdenas, C., Hernández, M., and Torres, J. (2014). An Exploratory Analysis of Heterogeneity on Regional Labour Markets and Unemployment Rates in Colombia: An MFACT approach. Borradores de Economía 802, Banco de la República de Colombia.
- Díaz, A. (2011). Spatial Unemployment Differentials in Colombia. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2011014, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).
- DNP (2012). Misión de Sistema de Ciudades. Algunos aspectos del análisis del Sistema de Ciudades. Technical report, Departamento Nacional de Planeación - DNP.

- Elhorst, J. Paul (2000). The Mystery Of Regional Unemployment Differentials: A Survey Of Theoretical And Empirical Explanations. ERSA conference papers ersa00p60, European Regional Science Association.
- Eslava, M., Meléndez, M., and Perry, G. (2014). Notas sobre Política de Desarrollo Productivo en Colombia. Documentos CEDE 012342, Universidad de los Andes -CEDE.
- Firpo, S. (2007). Efficient Semiparametric Estimation of Quantile Treatment Effects. *Econometrica*, 75(1):259–276.
- Fortin, N., Firpo, S., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):1001–44.
- Fortin, N., Firpo, S., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.
- Fortin, N., Firpo, S., and Lemieux, T. (2011). *Decomposition Methods in Economics*, volume 4 of *Handbook of Labor Economics*, chapter 1, pages 1–102. Elsevier.
- Galvis, L. (2002). Integración regional de los mercados laborales en Colombia, 1984-2000. Documentos de Trabajo sobre Economía Regional 002898, Banco de la República.
- Galvis, L. (2010). Comportamiento de los salarios reales en Colombia: Un análisis de convergencia condicional, 1984-2009. Documentos de Trabajo sobre Economía Regional 007047, Banco de la República.
- Galvis, L. (2013). Dinámica de crecimiento económico y demográfico regional en Colombia, 1985-2011. Documentos de Trabajo sobre Economía Regional 010791, Banco de la República.
- Galvis, L. and Meisel, A. (2012). Convergencia y trampas espaciales de pobreza en Colombia: Evidencia reciente. Documentos de Trabajo sobre Economía Regional 010287, Banco de la República.
- Galvis, L. and Meisel, A. (2016). Aspectos Regionales de la Movilidad Social y la Igualdad de Oportunidades en Colombia. *Revista de Economía del Rosario*, 17(2):257–297.

- García, A., Guataqui, J., and Rodríguez, M. (2009). Estimaciones de los determinantes de los ingresos laborales en Colombia con consideraciones diferenciales para asalariados y cuenta propia. Documentos de Trabajo 005756, Universidad del Rosario.
- García, A., Mesa, D. G., and Roa, M. (2008). Estructura salarial y segmentación en el mercado laboral de Colombia: un análisis de las siete principales ciudades, 2001-2005. Documentos de Trabajo 005105, Universidad del Rosario.
- Goraus, K., Van Der Velde, L., and Tyrowicz, J. (2013). What is the true gender wage gap? a comparative analysis using data from poland. Working Papers 2013-28, Faculty of Economic Sciences, University of Warsaw.
- Jaramillo, C., Nupia, O., and Romero, C. (2000). Integración En El Mercado Laboral Colombiano 1945-1998. Borradores de Economía 002896, Banco de la República.
- Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, 101(3):410–42.
- Londoño, E. (2013). Retornos a la Educación y Promedio Bayesiano de Modelos s partir de un pseudo panel: aplicación en Colombia. Documento de trabajo. mimeo, Universidad del Rosario.
- Machado, J. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4):445–465.
- Medina, C. and Saavedra, J. (2012). Formación para el Trabajo en Colombia. Borradores de Economía 740, Banco de la República de Colombia.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4):281–302.
- Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.
- Nupia, O. (1997). Integración espacial en los mercados laborales: evidencia para las regiones colombianas. *Revista Desarrollo y Sociedad*.

- Nupia, O. (2014). Distribución Regional de las Políticas de Desarrollo Productivo en Colombia y Brechas Regionales en Productividad y Empleo. Documentos CEDE 012059, Universidad de los Andes -CEDE.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709.
- Quiñonez, M. (2011). Diferencias regionales y capital humano: Examinando las brechas en los salarios de los individuos en Colombia. Tesis de maestría, Universidad del Valle.
- Romero, J. (2008). Diferencias sociales y regionales en el ingreso laboral de las principales ciudades colombianas, 2001-2004. *Revista de Economía del Rosario*.
- Taubman, P. and Wachter, M. (1987). Segmented labor markets. In Ashenfelter, O. and Layard, R., editors, *Handbook of Labor Economics*, volume 2 of *Handbook of Labor Economics*, chapter 21, pages 1183–1217. Elsevier.

A. Anexos

Figura 7: Distribución de ocupados por grupos en Bogotá y las 12 ciudades

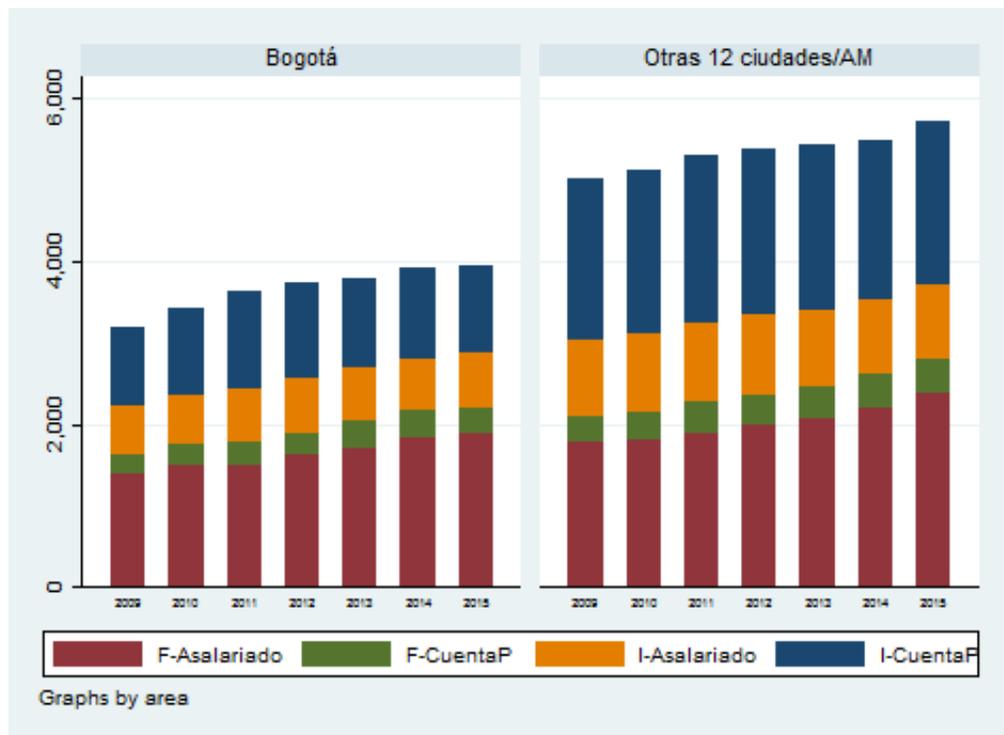


Figura 8: Ingresos laborales, Años de Educación y Edad por grupos de ocupados - Bogotá y las 12 ciudades.

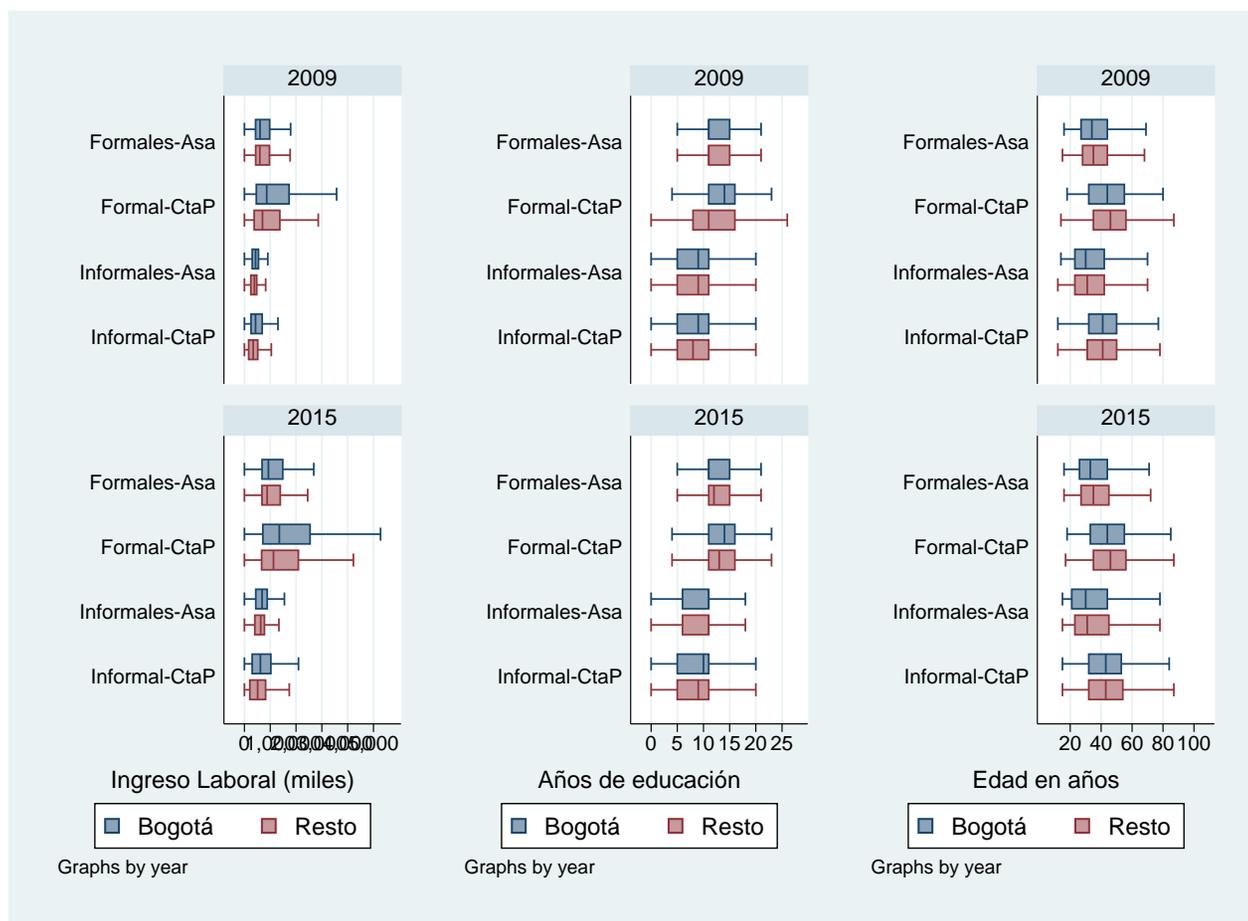


Figura 9: Distribución de ocupados por grupos - Bogotá y otras ciudades principales.

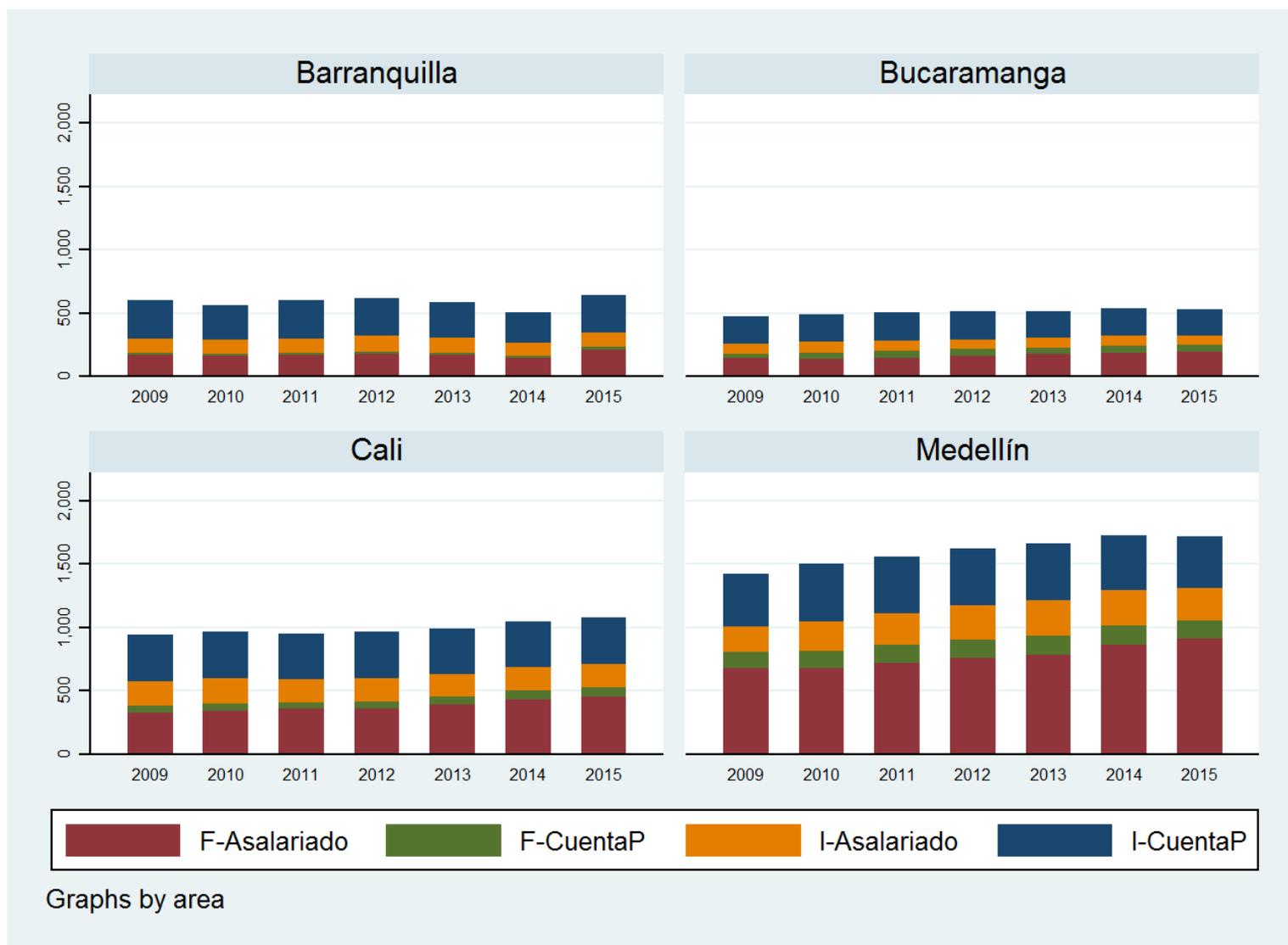


Figura 10: Ingresos y tamaño de los mercados laborales por grupos - Bogotá y otras ciudades principales.

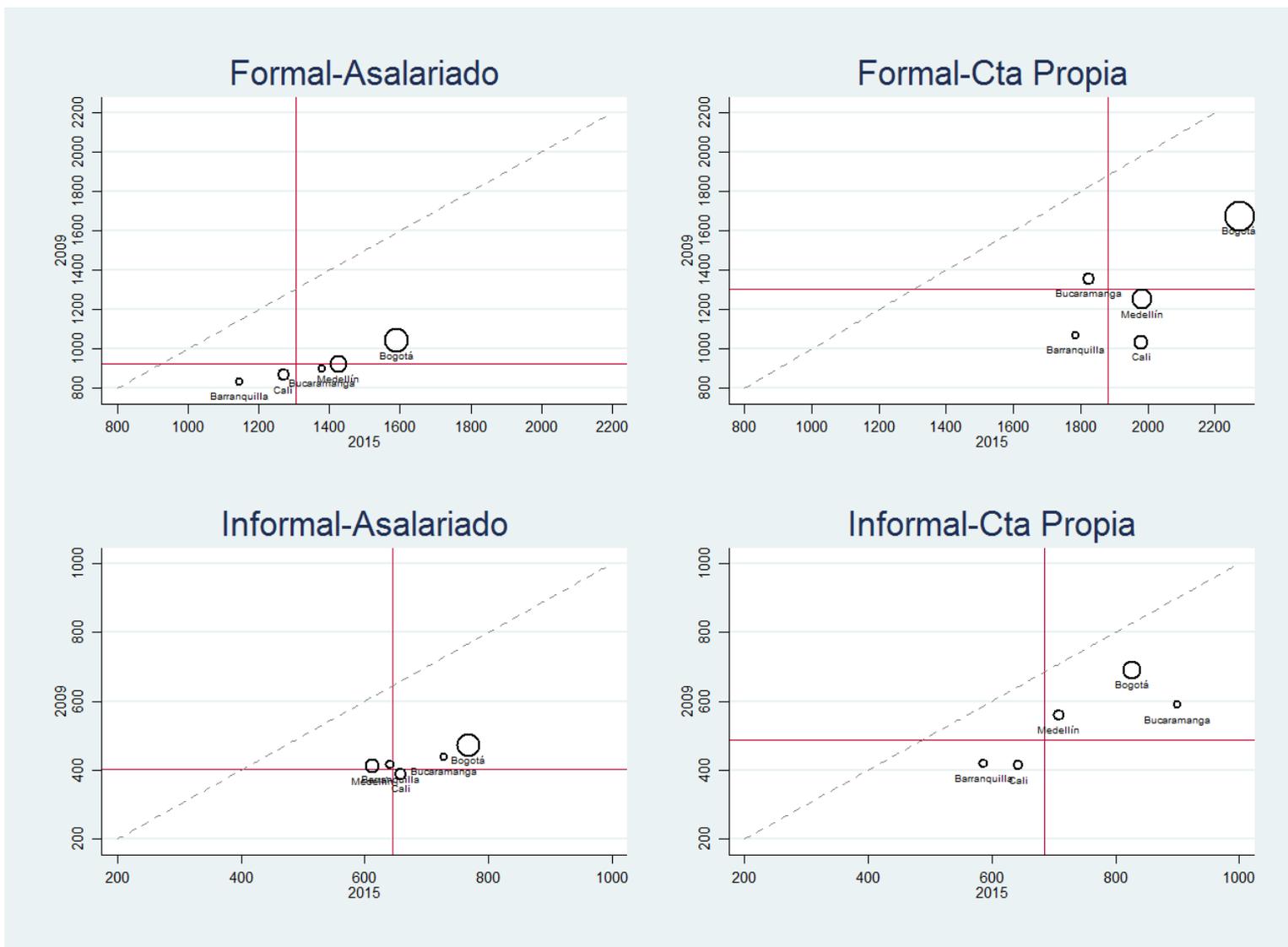


Figura 11: Ingresos laborales por grupos - Bogotá y otras ciudades principales.

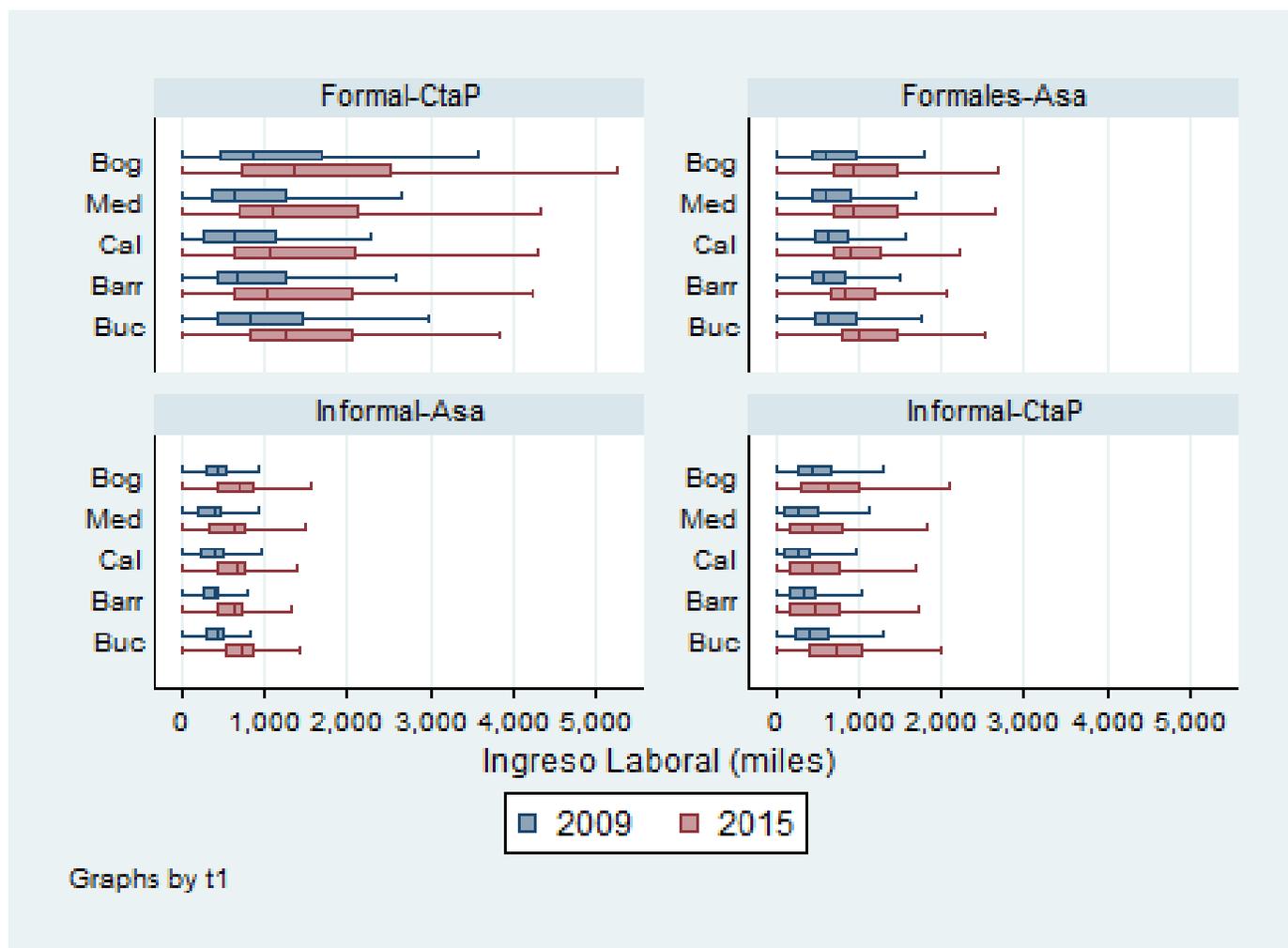


Figura 12: Educación y edad por grupos - Bogotá y otras ciudades principales.

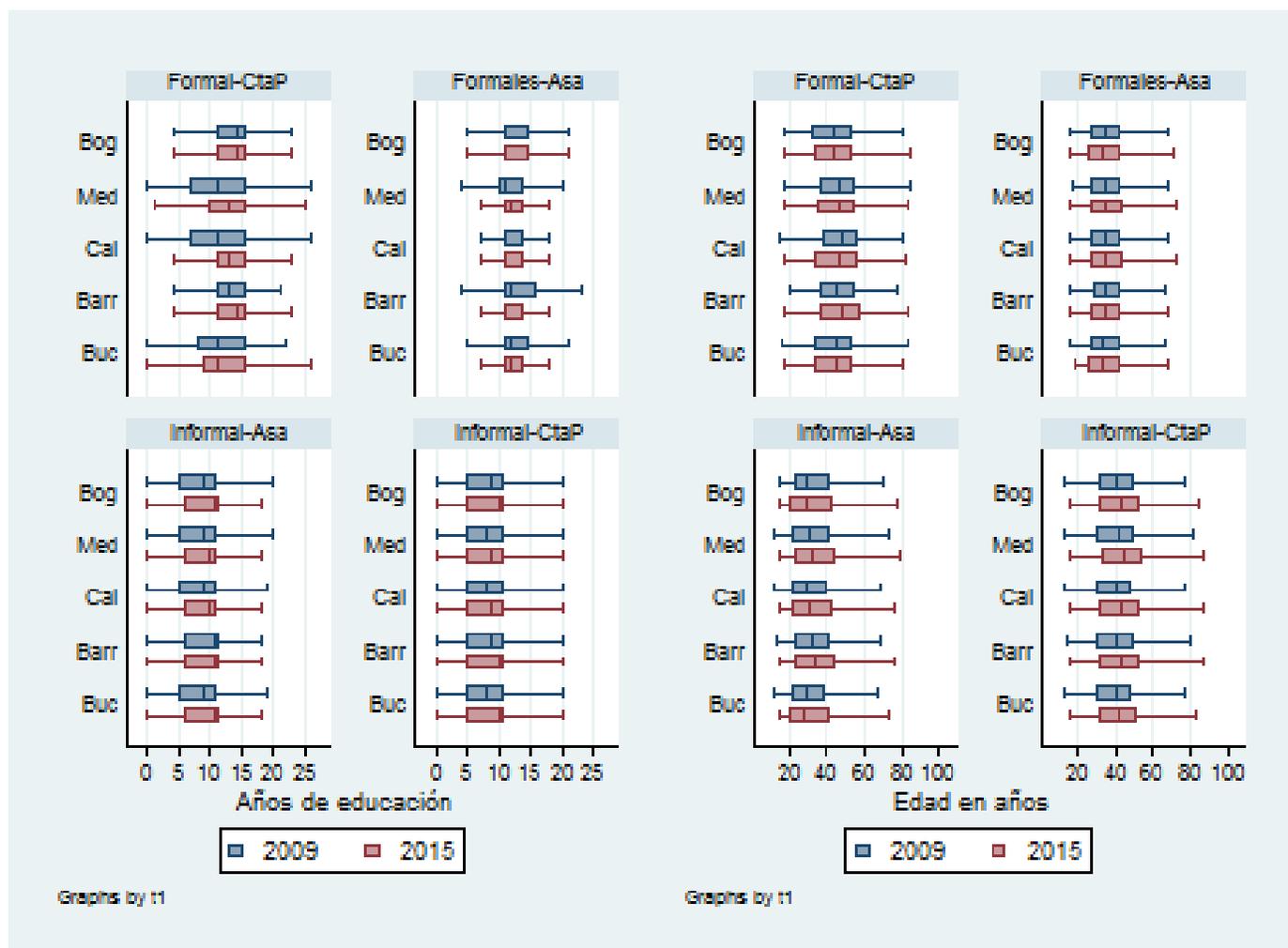


Figura 13: Diferencia en estructura (2009) - Bogotá y otras ciudades principales.

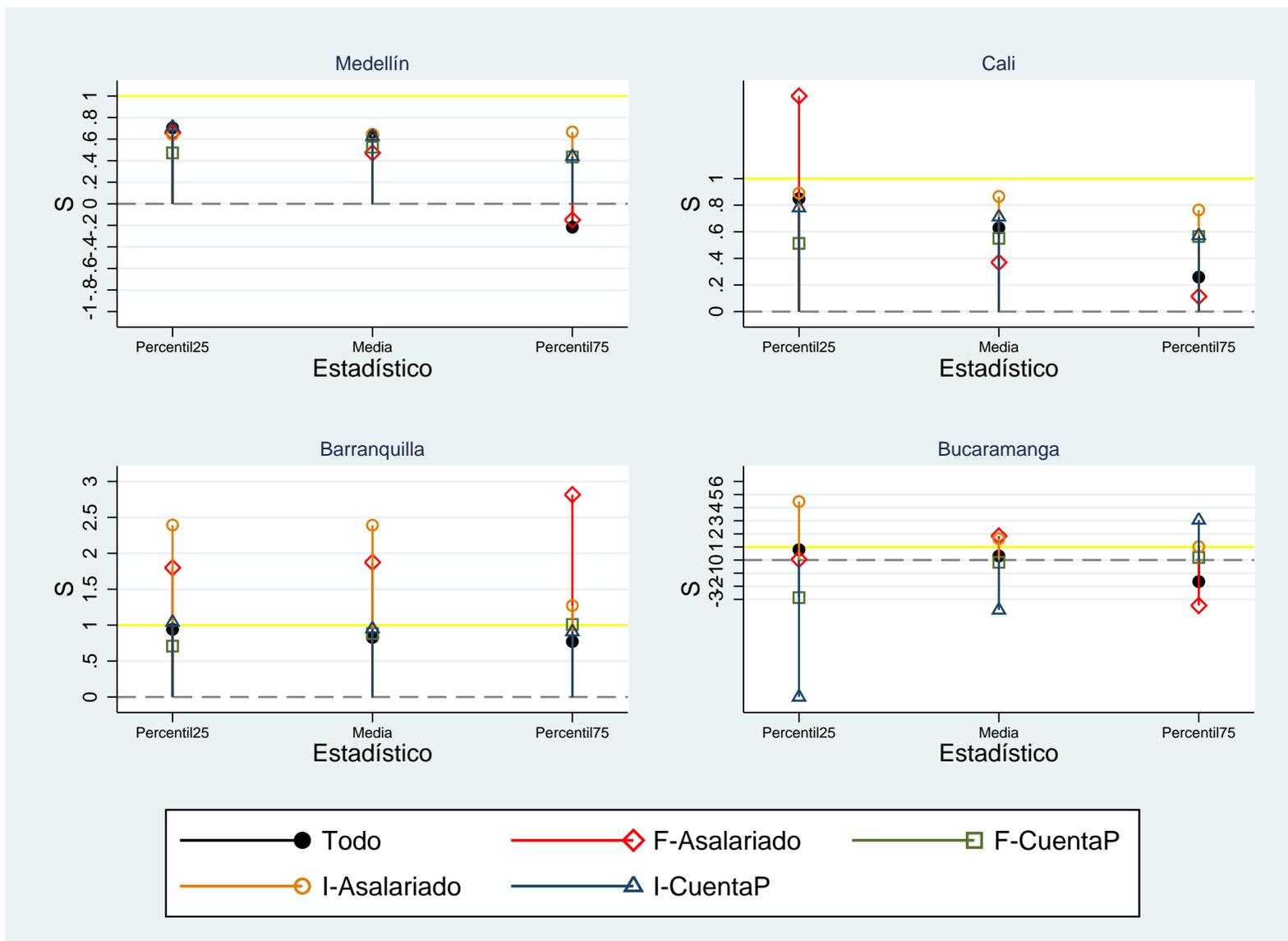


Figura 14: Diferencia en estructura (2010) - Bogotá y otras ciudades principales.

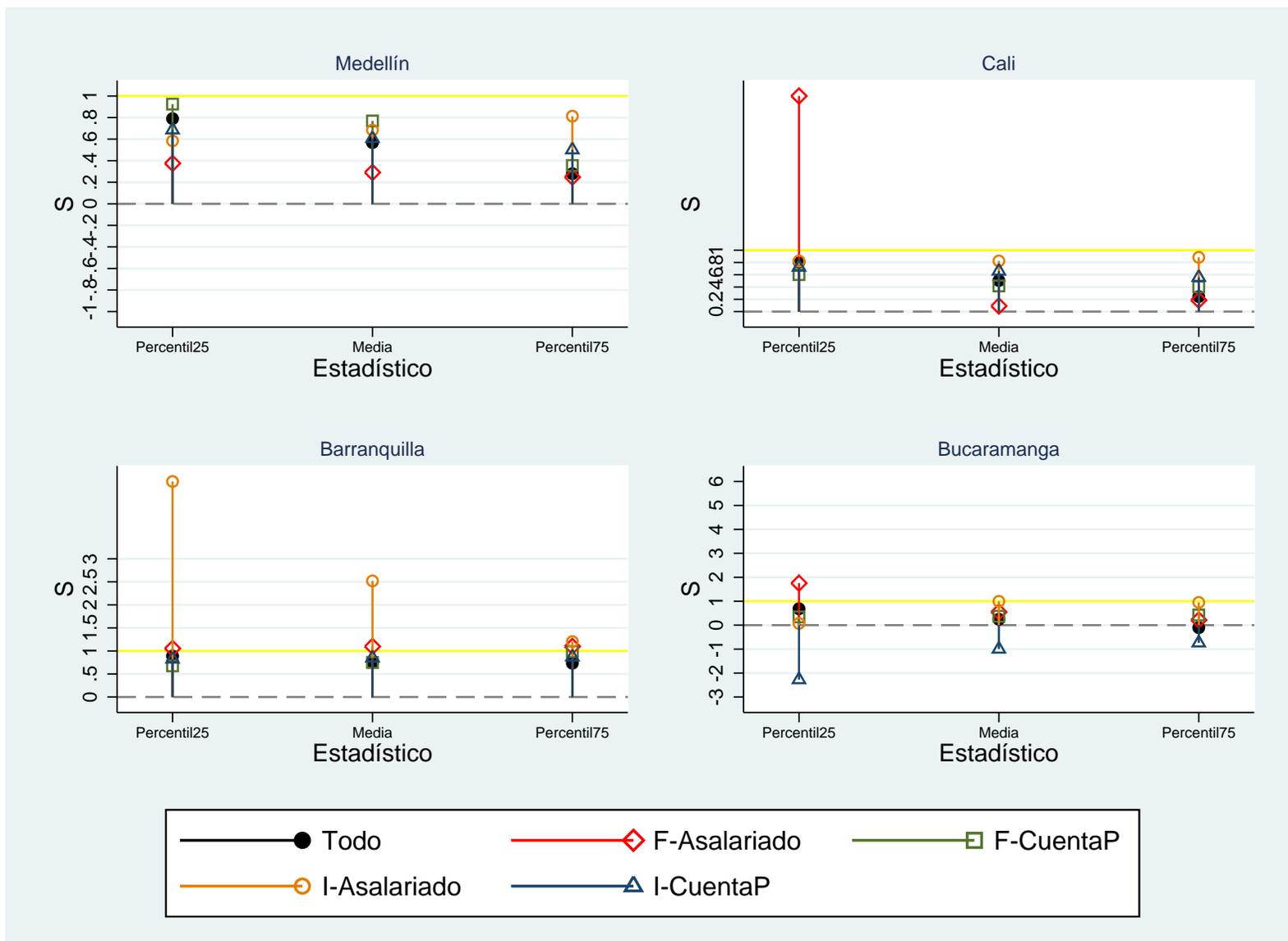


Figura 15: Diferencia en estructura (2011) - Bogotá y otras ciudades principales.

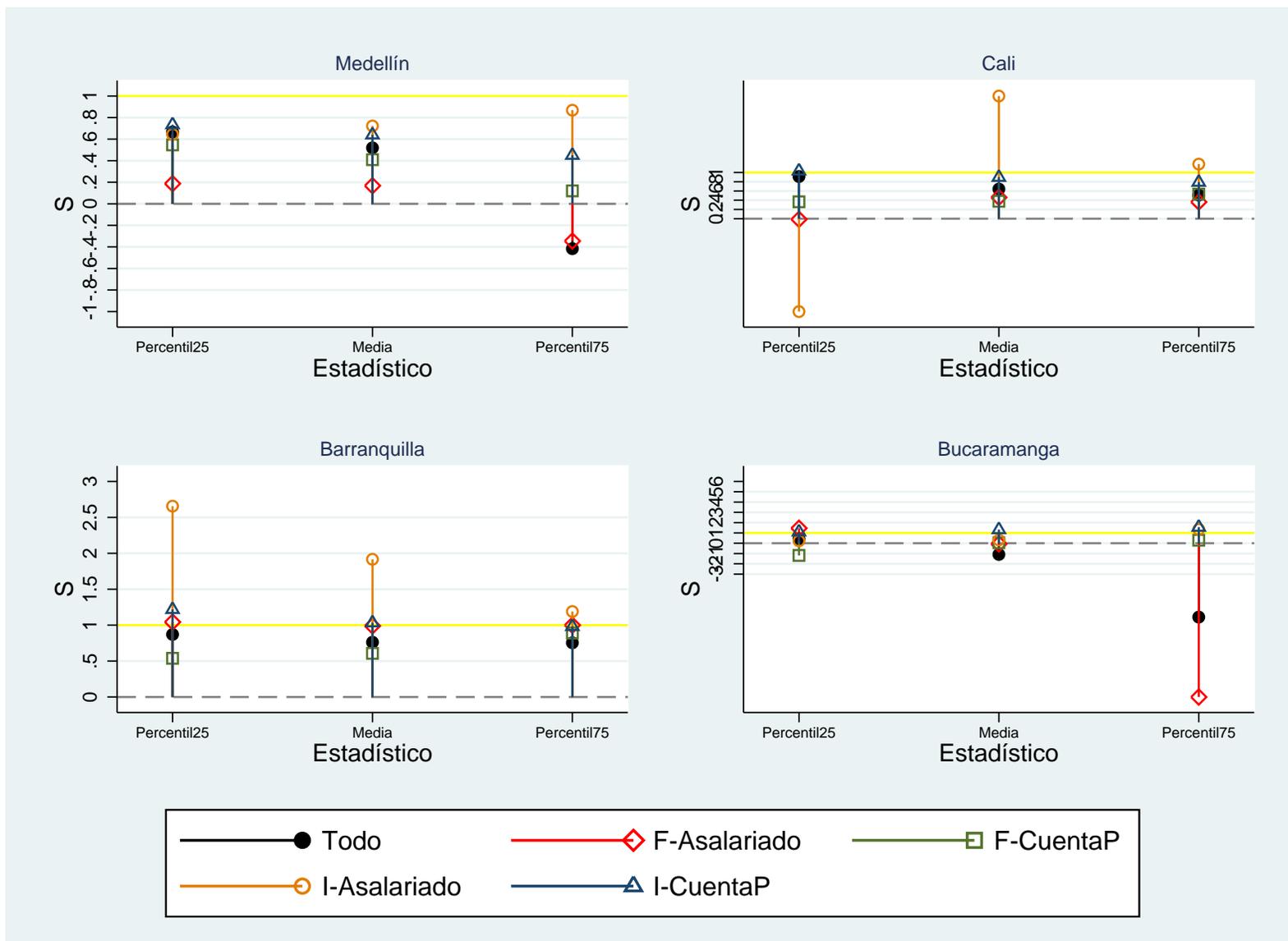


Figura 16: Diferencia en estructura (2012) - Bogotá y otras ciudades principales.

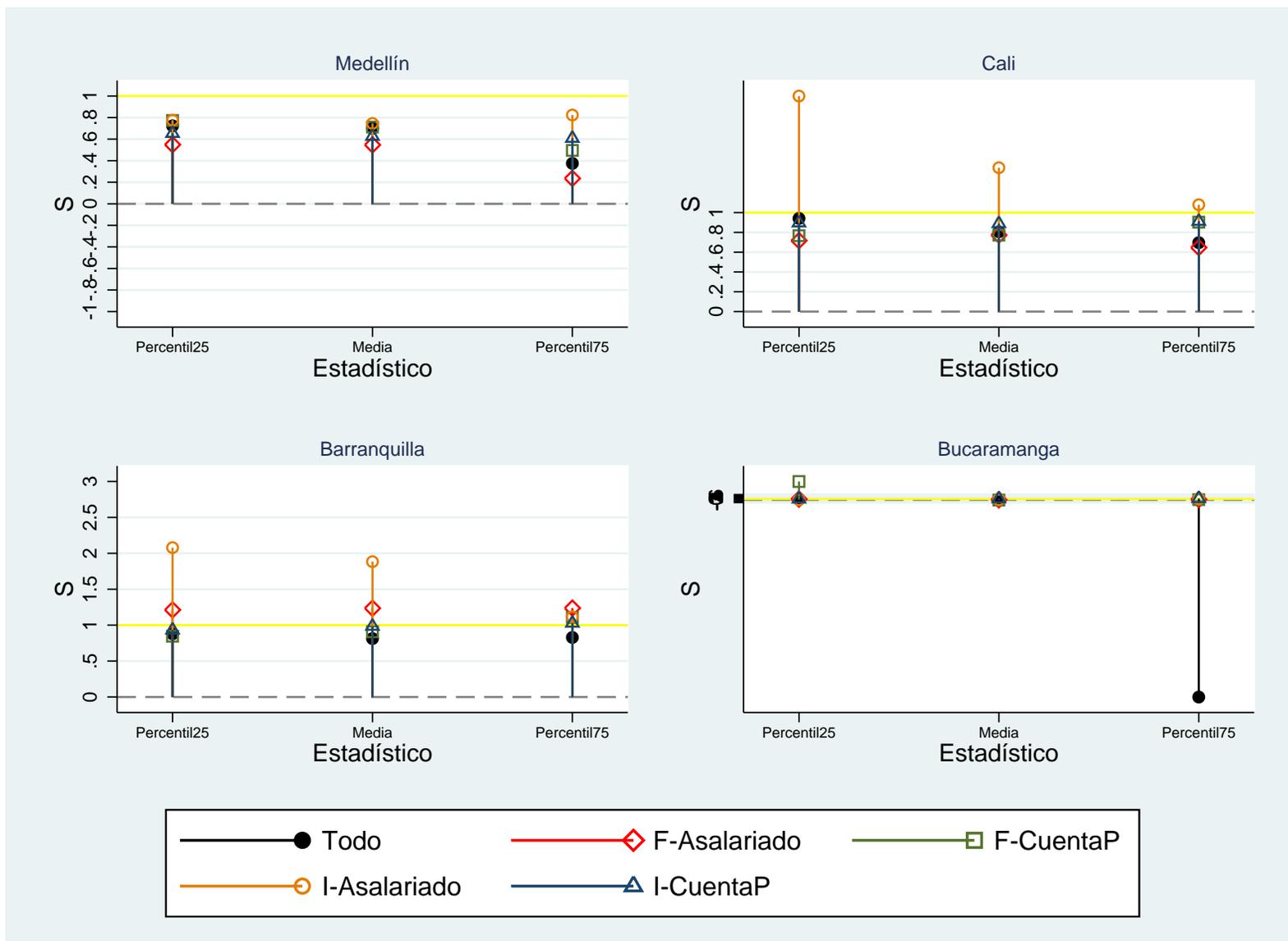


Figura 17: Diferencia en estructura (2013) - Bogotá y otras ciudades principales.

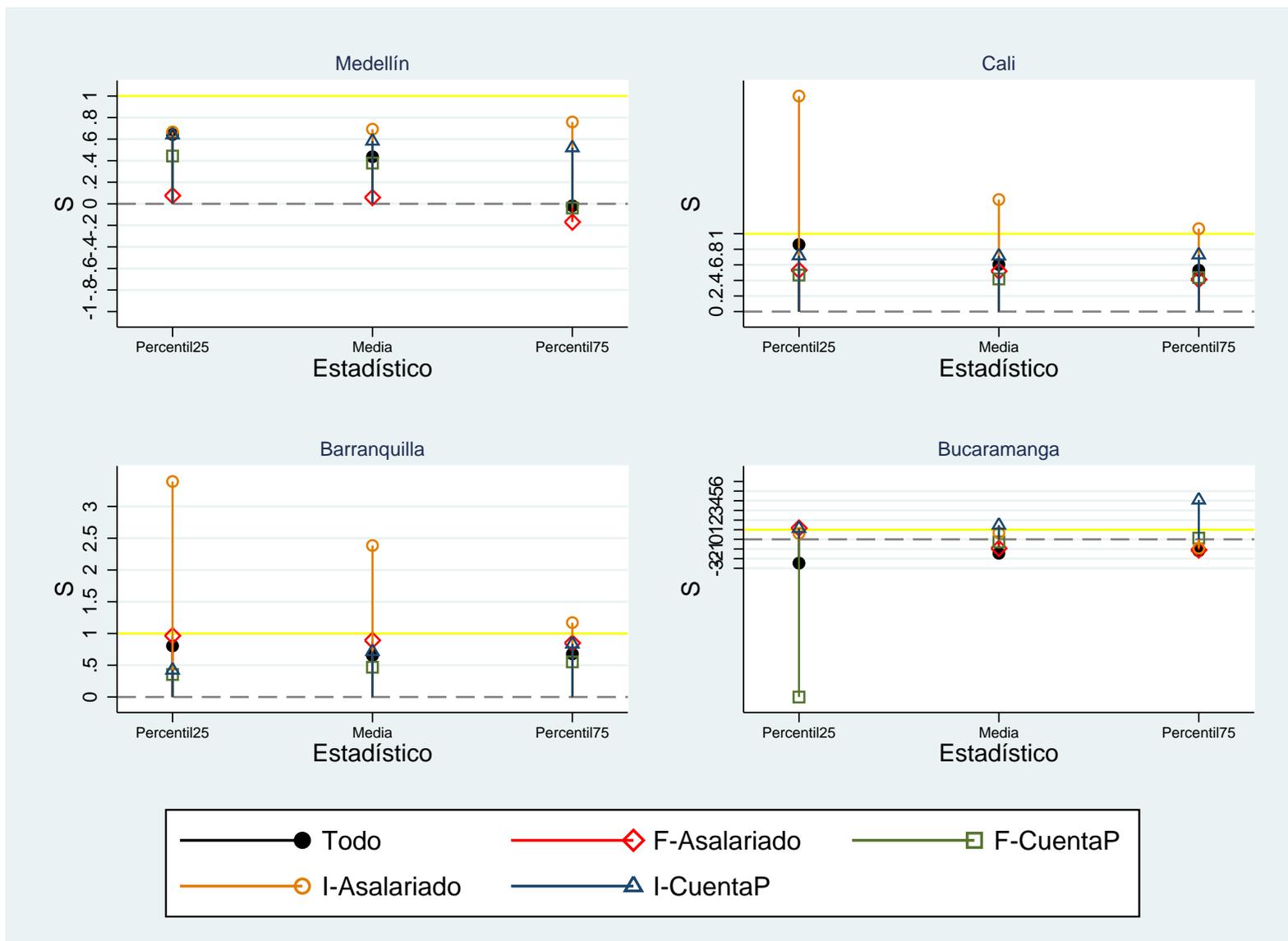


Figura 18: Diferencia en estructura (2014) - Bogotá y otras ciudades principales.

